

RECHERCHES  
LINGUISTIQUES  
DE VINCENNES

## Recherches linguistiques de Vincennes

32 | 2003

Grammaire et gabarits

---

# Les pluriels internes de l'arabe : système et conséquences pour l'architecture de la grammaire

Alain Kihm

---



### Édition électronique

URL : <http://journals.openedition.org/rlv/473>

DOI : 10.4000/rlv.473

ISSN : 1958-9239

### Éditeur

Presses universitaires de Vincennes

### Édition imprimée

Date de publication : 1 juin 2003

Pagination : 109-156

ISBN : 2-84292-137-2

ISSN : 0986-6124

### Référence électronique

Alain Kihm, « Les pluriels internes de l'arabe : système et conséquences pour l'architecture de la grammaire », *Recherches linguistiques de Vincennes* [En ligne], 32 | 2003, mis en ligne le 06 juin 2005, consulté le 19 avril 2019. URL : <http://journals.openedition.org/rlv/473> ; DOI : 10.4000/rlv.473

---

**Alain KIHM**

**CNRS – Laboratoire de Linguistique formelle – Université Paris 8**

**LES PLURIELS INTERNES DE L'ARABE :  
SYSTÈME ET CONSÉQUENCES POUR L'ARCHITECTURE  
DE LA GRAMMAIRE**

**RÉSUMÉ**

Les pluriels internes des noms et des adjectifs en arabe, aussi appelés pluriels « brisés », posent un défi à la théorie morphologique : ils ne mettent pas ouvertement en jeu l'affixation d'un morphème de pluralité, mais semblent plutôt reposer sur un contraste de formes qui, si elles partagent la même racine consonantique, diffèrent, entre autres, pour la vocalisation.

Cet article propose une analyse de la formation des pluriels internes qui dérive toutes les formes par un unique processus, à savoir l'insertion de l'un des trois glides /I/, /U/ et /A/ dans un *site* particulier localisé entre la 2<sup>e</sup> et la 3<sup>e</sup> consonne radicale, le site post-C<sub>2</sub>, avec pour effet la création d'une position d'attaque (C) ou de noyau (V).

Il explore ensuite les conséquences de cette analyse pour l'organisation générale de la grammaire. On y développe une version modifiée de la théorie dite « Morphologie Distribuée », telle que le lexique est conçu comme un ensemble d'ensembles d'éléments pourvus de significations mais dépourvus de forme, et l'épél est une fonction qui prend des ensembles d'éléments pour produire des objets phonologiques bien formés, à savoir des suites CV.

**MOTS-CLÉS**

CV (phonologie), ensemble (théorie des), fonctions, gabarit, glides, (non)concaténative (morphologie), site.

*La morphologie est la science qui traite des unités de son correspondant à une partie de l'idée et du groupement de ces unités [...]. Le vrai nom de la morphologie serait la théorie des signes et non des formes.* (Saussure 2002 : 182)

## 1. Introduction

Cet article est consacré aux pluriels internes des noms et des adjectifs arabes, qui, le nom l'indique, ne mettent pas en jeu une marque externe, affixale, de pluralité<sup>1</sup>. J'en développerai une analyse dont le thème central est le suivant : les pluriels internes (PI) s'opposent aux singuliers correspondants non par la présence d'un morphème de pluralité qui manquerait à ces derniers, mais par la seule différence entre deux états de la racine : l'état simple du singulier et l'état amplifié du PI. Autrement dit, les PI ne sont pas dérivés des singuliers correspondants, mais les deux formes représentent des réalisations distinctes, quoique connexes, d'une même racine. Dans cette mesure, la théorie ici défendue constitue une théorie non-morphémique (*a-morphous*) au sens d'Anderson (1992).

On verra toutefois que l'amplification de la racine, cause de la différence, ne se fait pas n'importe où, mais dans un *site* dont la position est parfaitement repérable. Je reprends cette notion de site à Guerssel & Lowenstamm (1990), qui l'utilisent avec succès pour analyser la dérivation verbale arabe, et aussi à Asfour (2001) qui, s'inspirant de ces auteurs, donne une analyse des PI proche de la mienne, quoique je m'en écarte sur quelques points importants. Un site est un lieu fixe où se produit une activité morphologique variable, en ce sens que ce n'est pas toujours la même représentation sémantique (pluralité, aspect, etc.) qui s'exprime dans et par le site. Par là, il est permis de considérer le site et l'activité qu'il accueille – les deux sont indissociables – comme un morphème, ainsi que le font Guerssel & Lowenstamm (1990). La différence entre ce type de morphème et le morphème "objet" serait alors que ce dernier peut varier dans sa forme (allomorphie) mais est constant dans son sémantisme, alors que le morphème-site ne change pas de forme, puisqu'il n'en a pas et consiste en une localisation, mais varie quant aux significations qui s'expriment à travers lui sous des formes phonologiques elles aussi changeantes. La nécessité que le morphème-site soit strictement localisable pour être fonctionnel – puisque seule la *position* du matériel phonologique signale qu'il remplit une fonction morphologique spécifique – explique qu'il se rencontre principalement (uniquement ?) dans les langues dont les lexèmes se conforment à des *gabarits* déterminés – ce pour quoi on peut aussi parler de morphème gabaritique pour cet objet morphologique<sup>2</sup>. Quant à savoir si une différence essentielle sépare cette conception du morphème-site de la conception non-morphémique d'Anderson (1992), c'est là une question sur laquelle je ne me prononce pas, car elle excède de beaucoup l'objet de cet article.

La présente théorie intègre par ailleurs de nombreux apports de la Morphologie distribuée (*Distributed Morphology* ou DM — *cf.*, par exemple,

Halle & Marantz 1993), en particulier l'hypothèse selon laquelle le lexique est un ensemble (d'ensembles) de traits abstraits ou éléments potentiellement pourvus d'une dénotation (les racines) ou d'un sens grammatical (les éléments fonctionnels); ou encore l'idée que la syntaxe est un cadre d'insertion pour les éléments lexicaux (lexèmes), qui forme ainsi l'interface entre le lexique et la morphologie, laquelle constitue la composante réalisationnelle ou d'épel de la grammaire, où les items du vocabulaire (les mots ou vocables) se voient finalement évoqués.

Je m'écarte néanmoins de la DM "classique" sur deux points. Premièrement, je fais l'hypothèse, que je m'efforcerai de démontrer, que nombre d'opérations que la DM confie à la syntaxe par le biais du déplacement des têtes peuvent s'accomplir au sein du lexique moyennant une formalisation ensembliste des processus lexicaux. On réalise ainsi une économie importante quant au nombre des têtes fonctionnelles tout en réduisant beaucoup le rôle du déplacement des têtes — dont il serait peut-être souhaitable de se dispenser tout à fait pour des raisons d'ordre général (*cf.* Chomsky 1999). En fait, il se pourrait que l'adjonction syntaxique à une tête fonctionnelle s'en trouvât limitée aux cas où un trait exprimé par un morphème n'est présent que parce que son porteur a été inséré dans une certaine structure syntaxique — on pense, par exemple, au cas structural. En outre, de façon plus immédiatement pertinente pour les buts de cette étude, un tel modèle permet de rendre compte très directement des particularités des PI en regard des pluriels affixaux que l'arabe connaît également.

Deuxièmement, la DM classique voit dans l'épel l'appariement d'un faisceau de traits (*feature bundle*) venu de la syntaxe avec une matrice de traits définissant un morphème-objet listé. Vu mes hypothèses à propos du lexique, j'y verrai plutôt l'application d'une fonction ou d'une composition de fonctions qui prend pour domaine un ensemble d'éléments et produit un objet morpho-phonologique dont l'analyse interne en morphèmes est un corrélât de l'application fonctionnelle (*cf.* Raffelsiefen 1992 et plus loin). Une telle conception est de toute façon nécessaire étant donné que mon explication des PI ne met justement pas en jeu de morphème-objet, et que le morphème-site dont elle fait usage ne s'accorde pas avec le caractère intrinsèquement concaténatif de l'appariement traits – morphèmes-objets. En outre, on le verra, cette conception révèle clairement en quoi consiste réellement la morphologie dite non concaténative. Au total, la théorie ici défendue peut être considérée comme relevant de la famille des théories morphologiques dites « inférentielles-réalisationnelles » (*cf.* Stump 2001).

Je commencerai par décrire rapidement le système des PI, afin d'en mettre en relief les caractères essentiels (section 2). Puis, je passerai tout aussi rapidement en revue quelques unes des théories qui ont été proposées pour rendre compte de ce système (section 3). A la section 4, je présenterai les éléments de ma propre théorie, que j'appliquerai ensuite à quelques cas typiques (section 5). Enfin, j'examinerai les conséquences qu'entraîne cette théorie particulière pour la théorie générale de la grammaire (section 6).

## 2. Description sommaire du système

En arabe, le pluriel des noms et des adjectifs se forme selon deux procédés : externe et interne, les pluriels « sains » et « brisés » des grammairiens arabes. Les pluriels externes (PE) sont suffixaux : /base + uuna/ pour les masculins et un petit nombre de féminins, /base + aatun/ pour les féminins et de nombreux masculins<sup>3</sup>. En gros, ils sont réservés aux items dont la forme et le sens manifestent l'opération d'une dérivation morphologique, ainsi du participe actif du verbe de mesure II *ʿallama* 'enseigner' : *muʿallimun* / *muʿallimatun* 'professeur, enseignant(e)' > *muʿallimuuna* / *muʿallimaatun* 'professeurs, enseignant(e)s' (voir plus loin sur cette notion de dérivation manifeste<sup>4</sup>). Tous les autres items forment des PI, qui constituent ainsi le cas non marqué ou par défaut, apparemment par modification des voyelles intercalées à la racine consonantique, avec éventuellement addition d'un préfixe /ʾa-/<sup>5</sup>. J'en donne ci-dessous quelques exemples qui en illustrent les caractères essentiels<sup>6</sup> :

- (1) KaLBun / KiLaaBun 'chien(s)'  
 NaJMun / NuJuuMun 'étoile(s)'  
 MaDiiNatun / MuDuNun 'ville(s)'  
 ĠuRFatun / ĠuRaFun 'chambre(s)'  
 ŠaMSun / ŠuMuuSun 'soleil(s)'  
 KiTaaBun / KuTuBun 'livre(s)'  
 NaJMun / ʾaNJuuMun 'étoile(s)'  
 JiSMun / ʾaJSaaMun 'corps'

Les caractères des PI qui se dégagent de (1) sont les suivants :

(a) **La multiplicité des gabarits.** Huit apparaissent en (1) : CiCaaC, CuCuuC, CuCuC, CuCaC, ʾaCCuC et ʾaCCaaC (C = toute consonne), mais ce n'est là qu'une petite partie des gabarits possibles, au nombre de 29 pour les racines bi- et triconsonantiques (2C et 3C – cf. Wright 1896-1898/1991 ; Blachère & Gaudefroy-Demombynes 1975), à quoi s'ajoute le gabarit unique CaCaaCi (i) C des quadriconsonantiques et plus (4C +)<sup>7</sup>. Certes, la fréquence de ces 29 gabarits n'est pas uniforme, certains sont très bien représentés, d'autres sont rares, mais tous sont attestés.

(b) **L'imprévisibilité du gabarit du PI étant donné celui du singulier.** Des corrélations existent, mais elles sont partielles et non systématiques : les études les plus fouillées (cf. Levy 1971 ; Ratcliffe 1998) montrent que la probabilité pour qu'un singulier d'un certain type se pluralise selon tel gabarit dépasse rarement 50 %. Dans ces conditions, tout se passe comme si les appariements singulier-PI devaient être mémorisés un à un, sans que les apprenants puissent jamais espérer une généralisation certaine.

(c) **La non-pertinence du genre du singulier :** le féminin marqué *madiinatun* 'ville' a pour PI *mudunun* de même gabarit que le masculin *kitaabun* 'livre', dont le PI est *kutubun* ; *šamsun* 'soleil', féminin « par convention », a pour PI *šumuusun*, comme le masculin *najmun* 'étoile' a

*nujuumun*<sup>8</sup>. Les PI se distinguent ainsi des PE, pour lesquels le genre du singulier détermine le choix du suffixe pluralisateur.

(d) **L'existence fréquente de variantes** associées à un même singulier, comme le montre l'exemple de *najmun* qui a les deux PI *nujuumun* et *'anjmun*. Ces variantes – il peut y en avoir plus de deux – sont sémantiquement équivalentes et ne se distinguent, éventuellement, que sur le plan stylistique.

(e) **La productivité**: tout néologisme permettant l'extraction d'une suite de trois consonnes est susceptible de recevoir un PI, témoin l'emprunt récent *filmun* / *'aflaamun* 'film(s)' (pseudo racine *flm*).

(f) **La possibilité de (re) pluraliser un PI**, en formant ou bien un nouveau PI ou bien un PE: cf. *yadun* / *'aydin* / *'ayaadin* 'main(s)', *Harbun* / *Huruubun* 'guerre(s)' / *Huruubaatun* 'campagnes militaires'. La repluralisation n'a pas, semble-t-il, d'effet sémantique dans le premier exemple; elle en a un, en revanche, dans le second: «campagnes» se comprend comme «suite de guerres». Il importe de noter qu'il est tout à fait impossible de repluraliser ainsi un PE (\**Huruubaataatun*/).

(g) **La non spécialisation du procédé**: les diminutifs se forment en imposant à la racine un gabarit CuCayC pour les 3C, CuCayCiC pour les 4C +, semblable aux gabarits des PI: cf. *kalbun* 'chien' / *kulaybun* 'petit chien' (comparer *kilaabun* 'chiens'), *'aqrabun* 'scorpion' / *'uqayribun* 'petit scorpion' (comparer *'aqaaribu* 'scorpions'). La formation du nom verbal ou *maSdar* met aussi en jeu ce procédé (cf. Kihm 2003).

Tous ces caractères soulignent la nature *dérivationnelle* des PI<sup>9</sup>. La multiplicité et l'imprévisibilité sont notoirement le propre des formations dérivationnelles (cf. *nettoyage*, non pas \*/nettoisement/ vs *voussoiement*, non pas \*/voussoyage/) <sup>10</sup>. Un morphème dérivationnel peut s'attacher à un lexème d'un genre quelconque et imposer un genre associé à la dérivation elle-même (cf. les diminutifs allemands en *-chen*, toujours neutres). L'existence de variantes, sans être très fréquente, n'est pas rare (cf. *enfouissage* ou *enfouissement*)<sup>11</sup>. La productivité conditionnelle, liée à l'identité du morphème et/ou à telles propriétés de la base, est un trait distinctif des procédés dérivationnels, qui les oppose aux flexions, le plus souvent productives sans restriction. Les morphèmes dérivationnels sont souvent polyvalents quant à leurs corrélats sémantiques: *-eur* en français dérive des noms d'agent (*voleur*), mais aussi des noms d'instruments (*broyeur*). En fait, le seul caractère apparemment déviant des PI, en tant que dérivation, est (f), pour autant que les opérations dérivationnelles ne se superposent pas, sauf à modifier à chaque fois la catégorie de l'item dérivé (cf. *urbain* > *urbaniser* > *urbanisation*, mais \*/urbanisationneur/ vs *urbanisateur*). Mais, d'une part, les "doubles pluriels" sont attestés dans plus d'une langue, y compris lorsque le pluriel est affixal (cf. Stump 1990 à propos du breton); d'autre part, il n'est pas inconcevable que la pluralisation d'un PI (d'un pluriel en général) implique la neutralisation du trait de pluralité de celui-ci, donc de l'opération dérivationnelle (nous en verrons plus loin un analogue). S'il en est ainsi, (f) ne s'oppose pas à la dérivationnalité des PI<sup>12</sup>.

Celle-ci explique alors une propriété centrale des PI signalée tout à l'heure, à savoir leur incompatibilité avec les items manifestement dérivés. Il importe de s'arrêter sur cette notion de « dérivation manifeste », car elle n'est pas sans conséquence.

Fondamentalement, il s'agit d'une relation sémantique. J'ai déjà donné l'exemple de *mu'allimun* 'professeur', participe actif de *'allama* 'enseigner', mesure II, de sens causatif (cf. *'alima* 'savoir') ; j'aurais pu citer tout aussi bien un « nom de profession » comme *xayyaaTun* 'tailleur' formé par application du gabarit CaCCaaC à la racine *x (y) T* (cf. *xaaTa* 'coudre'). Vu la transparence de la relation entre la signification du dérivé et le sens de la racine – un professeur est un enseignant, le tailleur est celui qui coud (ou taille) les vêtements – il est permis de supposer que les locuteurs ont conscience de la connexion dérivationnelle<sup>13</sup>. On peut en tout cas faire l'hypothèse, plus intéressante pour nous, que cette connexion est spécifiquement représentée dans la grammaire intériorisée par lesdits locuteurs. On sait que les items de cette sorte ne connaissent que le PE (*mu'allimuuna*, *xayyaaTuuna*).

Soit d'autre part un item comme *maktabun* 'bureau' : c'est également un dérivé, un « nom de lieu et de temps » (*ismu l-makaani wa-z-zamaani*) formé par application à la racine *ktb* (cf. *kataba* 'écrire') du gabarit maCCa/iC (cf. aussi *majlisun* 'conseil, assemblée' à côté de *jalasa* 's'asseoir', *masjidun* 'mosquée' à côté de *sajada* 'se prosterner', *manzilun* 'maison, demeure' à côté de *nazala* 'descendre', etc.). On voit toutefois qu'ici, la relation sémantique est bien moins transparente : un bureau ne se définit pas simplement comme un lieu où l'on écrit, et une mosquée ou une assemblée sont bien autre chose qu'un endroit où les gens se prosternent ou une occasion pour eux de s'asseoir. Et il y a aussi les noms d'instruments (*ismu l-'aalati*) tels que *miftaaHun* 'clef' (cf. *fataHa* 'ouvrir'), *mirwaHatun* 'éventail' (cf. *raaHa* 'souffler'), ou encore *miqaSSun* 'ciseaux' (cf. *qaSSa* 'couper'). Là encore, le lien entre, par exemple, « clef » et « ouvrir » n'est ni immédiat ni univoque.

Il apparaît ainsi que ces items, quoique formellement dérivés, et se distinguant par là de noms primitifs tels que *kalbun* 'chien' ou *jabalun* 'montagne' qui ne consistent en rien d'autre qu'une racine et une mélodie vocalique, constituent des items indépendants pour lesquels il n'existe pas dans la grammaire ni, probablement, dans l'esprit des locuteurs de connexion spécifique avec les autres items qui partagent la même racine (de la même façon qu'un francophone n'est sans doute pas conscient du lien entre « éventail » et « vent »). On parle alors de « relexicalisation », voulant dire par là que l'opération dérivationnelle est pour ainsi dire annulée. Le signe de cette annulation est que les items relexicalisés ne forment que des PI, et cela selon le gabarit des 4C + (cf. *makaatibu*, *mafaatiiHu*, etc.), ce qui montre que le /m/ initial, bien qu'à l'origine morphème extérieur à la racine, s'est vu réanalysé comme une consonne radicale<sup>14</sup>.

Selon Ratcliffe (1998), cet état de fait indique que, pour ce qui est de la morphologie nominale, le lexique de l'arabe se divise entre deux niveaux, I et II. Le niveau I comprend les items non soumis à des opérations morphologiques de

nature dérivationnelle ou pour lesquels lesdites opérations se sont vues annulées pour des raisons sémantiques. Le niveau II comprend les items résultant de processus dérivationnels non annulés. Seuls les items du niveau I sont accessibles à des opérations morphologiques internes à la racine, c'est-à-dire la formation d'un PI ou d'un diminutif, tandis que ceux du niveau II leur sont opaques et n'acceptent que des opérations à la limite externe de la racine<sup>15</sup>.

Cette division, outre qu'elle ne convient pas à la conception du lexique qui sera développée plus avant, paraît inutile dès lors qu'on a établi la nature dérivationnelle de la formation des PI. En effet, l'impossibilité de former un PI d'un item visiblement dérivé découle alors directement du principe général déjà évoqué selon lequel une dérivation qui ne change pas la catégorie syntaxique de l'item auquel elle s'applique ne peut se surajouter à une dérivation déjà accomplie (cf. la malformation de \*/urbanisationneur/).

On serait alors amené à considérer les PE comme étant de nature flexionnelle, en dépit d'une généralisation – la nature globalement dérivationnelle de la pluralisation – à laquelle on répugne pourtant à renoncer, car elle a en sa faveur de solides arguments (cf. Beard 1995). Il y a là un problème qui laisse à penser que la raison de la répartition des PI et des PE est à chercher ailleurs que dans la distinction, de toute façon fragile, de la dérivation et de la flexion, sans pour autant remettre en cause l'explication à laquelle nous sommes parvenus. J'y reviendrai.

Quoi qu'il en soit, on remarquera que l'analyse qui précède nous donne aussi l'explication des doubles pluriels du type '*aydin* – '*ayaadin* 'mains' et '*Huruubun* 'guerre(s)' – '*Huruubaatun* 'campagnes militaires'. Le premier est simplement un cas d'annulation de l'opération morphologique, si bien que '*aydin* se trouve dépluralisé au moment où l'opération produisant '*ayaadin* s'applique, tout comme '*miftaaHun* est « dé-dérivé » au même point<sup>16</sup>. La synonymie complète des deux formes s'ensuit. Pour le second, en revanche, l'opération n'est pas annulée, '*Huruubun* reste pluriel, si bien que sa pluralisation, nécessairement externe, ne peut rester sans effet.

Enfin, il importe d'énoncer un principe fondamental de la morphologie arabe, dont le domaine de validité excède la formation des PI, mais qui la régit tout particulièrement. Ce principe se formule ainsi : toute forme morphologiquement traitée doit reposer sur une racine au moins tri- et au plus quadriconsonantique. En d'autres termes, le triconsonantisme des racines arabes et, plus généralement, sémitiques, souvent vu comme un trait essentiel, est en fait une propriété morphologique (cf. Bohas 1997). Quant aux PI, le principe a pour conséquence que les racines 1C et 2C s'ajoutent une ou deux consonnes « par défaut » pour les former :

- (2) a. Šaa'un / ŠiYaaHun ~ ŠiWaaH<sup>um</sup> 'mouton(s)'  
 b. DaMun / DiMaa'un 'sang(s)'

On considère la racine comme étant 1C en (2a) parce que /š/ est la seule consonne que le singulier et le pluriel ont en commun<sup>17</sup>. Les deux consonnes par défaut



ajoutées pour parvenir au triconsonantisme requis par la morphologie sont /y/ ou /w/ en C<sub>2</sub> et /h/ en C<sub>3</sub>, c'est-à-dire, on le note pour plus tard, des segments du type «glide». En (2b), la consonne ajoutée aux deux de la racine *dm* est l'occlusive glottale, également du type «glide» (voir plus loin). Pour les racines 5C, le même principe entraîne en général la chute de la cinquième consonne au PI<sup>18</sup>:

- (3) <sup>c</sup>aNKaBuuTun / <sup>c</sup>aNaaKiBu 'araignée(s)'

On se gardera de voir là un exemple de morphologie soustractive (cf. Anderson 1992). On y verra plutôt l'illustration du caractère procustéen de la taille de la racine exigée par les opérations de la morphologie.

### 3. Hypothèses existantes sur la formation des PI

La tradition d'étude des PI est millénaire. Elle prend sa source dans les travaux des grammairiens arabes dont je ne dirai rien, faute de l'érudition nécessaire (cf. Bohas & Guillaume 1984). Les linguistes sémitisants du XIX<sup>e</sup> siècle se sont surtout attachés à rechercher l'origine de ces pluriels atypiques (cf. Derenbourg 1867; Barth 1904; et Corriente 1971 pour une bibliographie commentée)<sup>19</sup>. Je m'autoriserai donc à ne considérer que les travaux de grammaire générative, dont les plus importants à ce jour (et à ma connaissance) sont Levy (1971), McCarthy (1982), McCarthy & Prince (1990), Ratcliffe (1998) et Asfour (2001)<sup>20</sup>. Mon commentaire dans cette section s'attachera uniquement à McCarthy & Prince (1990) et à Ratcliffe (1998)<sup>21</sup>.

Ce n'est pas faire injure à ces auteurs que d'affirmer qu'ils cherchent avant tout à «normaliser» les PI en s'efforçant de montrer que ceux-ci mettent en jeu l'analogue d'un *morphème* pluriel. McCarthy & Prince (1990) isolent un schème «dominant» qui consiste en un pied iambique (brève – longue) sur une partie «prosodiquement circonscrite» du squelette CV... auquel se lient les segments. Cette partie correspond à la première syllabe lourde, longue, bimoraïque, soit /kvl/ dans *kalbun* 'chien', /nvj/ dans *najmun* 'étoile', /jvb/ dans *jabalun* 'montagne', /jvn/ dans *jundubun* 'sauterelle', etc., où *v* désigne une position vocalique<sup>22</sup>. La formation d'un PI consiste à remplacer cette syllabe par un pied iambique, c'est-à-dire la succession d'une syllabe brève et d'une syllabe longue, d'où //kilaa/bun/, //nujuu/mun/, //jibaa/lun/, //janaa/dibun/, selon la formule /CVC.../ > /CVCVV.../. La valeur de la voyelle brève, /i/, /u/ ou /a/, et celle de la longue, /aa/ ou /uu/, très rarement /ii/, sont, selon McCarthy & Prince, idiosyncrasiques et sans pertinence quant à l'opération elle-même.

Pour ces auteurs, ce remplacement équivaut à une affixation, si ce n'est que la suite phonologique à laquelle on ajoute n'est pas égale au lexème entier (comme dans l'anglais *dogs* /dog-Z/ 'chiens'), mais à une partie prosodiquement définie du squelette CV que le lexème instancie. Autrement dit, une langue comme l'anglais généralise la pluralisation sur les lexèmes selon la formule /LEX-Z/, où LEX égale tous les lexèmes existants et possibles, sauf une poignée d'exceptions (*foot* / *feet*, etc.), alors que l'arabe la généralise sur les squelettes.

Une autre différence est que l'affixe ou morphème mis en jeu par les PI arabes n'est pas un segment ou une suite continue de segments, mais deux mores auxquelles les segments se lient selon les principes de l'autosegmentalité<sup>23</sup>. Un tel morphème prosodique n'en est pas moins un morphème-objet, dont les allomorphes sont les diverses voyelles qui matérialisent le pied iambique, et qui se concatène à une base de manière un peu particulière. Tout cela reste bien différent du morphème-site dont je vais défendre l'hypothèse.

Il y a bien des choses à reprendre à ces thèses. On peut faire remarquer, par exemple, que la prédominance de la formule iambique n'est pas si évidente. Les auteurs sont forcés de reconnaître l'existence d'au moins une autre prosodie associée aux PI, qu'ils nomment "trochaïque" car elle consiste en une suite de deux syllabes brèves (cf. *kutubun* 'livres', *ğurafun* 'chambres', etc.) ou d'une syllabe longue et d'une brève (cf. *'anjum*<sup>um</sup> 'étoiles', etc.), et sur laquelle ils ne s'étendent guère<sup>24</sup>. Le nombre de ces PI, et d'autres qui, sans être trochaïques, ne satisfont pas non plus à la formule iambique, est trop important pour qu'on puisse les ravalier au statut d'exceptions, comme on s'autorise à le faire avec les pluriels apophoniques de l'anglais. Cela signifie qu'il existe au moins un autre morphème prosodique de pluriel en arabe, non marginal, et dont l'alternance avec le morphème iambique n'est pas morphologiquement ou phonologiquement prévisible, à la différence des divers morphèmes de pluriel de l'allemand, par exemple.

Cette absence de bi-univocité du trait pluriel vers le morphème le réalisant se vérifie aussi dans l'autre sens. Je veux dire par là que l'iambicité n'est pas univoquement associée à la pluralité: *kilaabun* 'chiens' est pluriel, mais *kitaabun* 'livre' est singulier<sup>25</sup>.

L'analyse de McCarthy & Prince ne dit rien non plus de formations telles que *mayyitun* / *mawtaa* 'mort(s)', *taajirun* / *tujjaarun* 'commerçant(s)', ou *diibaajun* / *dabaabiju* 'brocart(s)', qui, tout en contenant un pied iambique, ou bien ne l'ont pas dans la position attendue (*mawtaa*), ou bien témoignent d'autres processus – gémiation ou duplication de C<sub>2</sub> – qui paraissent tout autant constitutifs du PI, alors même que leur rapport avec l'iambisation ne saute pas aux yeux.

Tout cela ne laisse pas de jeter le doute sur la morphémicité (au sens de morphème-objet) attribuée à la prosodie iambique, plus précisément à l'insertion bimoraïque, car l'autre composante du pied, la voyelle brève entre C<sub>1</sub> et C<sub>2</sub> est évidemment épenthétique (cf. Ratcliffe 1998 : 80). De même, le préfixe /a/ qui apparaît dans *'ajsaamun* ne peut être considéré comme un morphème de pluralité, pour des raisons dans lesquelles je ne peux entrer ici, mais que l'on trouvera détaillées chez Ratcliffe.

McCarthy & Prince (1990) voient dans cette insertion un morphème affixé, plus précisément infixé, qu'ils comparent à des éléments qu'ils estiment semblables dans diverses langues: tagalog *-um-*, marque de focus de l'Agent, infixé après la première consonne (cf. *tawag* 'appeler' > *tumawag*); reduplication du premier pied dissyllabique en yidin<sup>y</sup> (Australie) pour former des pluriels (cf.

*mulari* > *mula-mulari* ‘initié(s)’, *gindalba* > *gindal-gindalba* ‘lézard(s)’ ; et d’autres encore.

En fait, la question n’est pas tant de savoir si l’insertion bimoraïque est ou non un affixe, car cela pourrait n’être qu’une pure affaire de terminologie. Elle est plutôt de savoir si le matériel ainsi affixé ou ajouté peut s’analyser ou non comme un morphème-objet, à savoir l’association régulière d’une forme ou d’un ensemble de formes et d’une signification. Avec cette définition, le *-um-* du tagalog est sans conteste un morphème, qui n’a de particulier que le lieu de son attachement. Les formes diverses des pluriels yidin<sup>9</sup>, pour leur part, se laissent prédire à coup sûr, puisqu’elles sont les copies exactes du pied qui suit immédiatement<sup>26</sup>. Ce n’est donc pas le polymorphisme qui pose un soi un problème, mais bien sa prévisibilité.

Afin d’ordonner l’étude des PI, il est commode de répartir les noms et les adjectifs de l’arabe en groupes selon le nombre de leurs consonnes radicales (d’une à cinq) et selon leur genre (masculin ou féminin). Soit le Groupe 1, qui comprend les noms masculins non dérivés dont la racine comporte trois consonnes ou moins. C’est dans ce groupe que les noms se pluralisent le plus par insertion bimoraïque, laquelle est assez bien prédictible, puisqu’elle se réalise /aa/, c’est-à-dire /a/ lié aux deux positions, pour 36 % des singuliers CaCCun (cf. *kalbun* / *kilaabun*), 54 % des singuliers CiCCun (cf. *jismun* > ‘*ajsaamun*) et 58 % des singuliers CuCCun (cf. *quflun* / ‘*aqfaalun* ‘verrou(s)’, variante de *qufuulun*). Qui plus est, la grande majorité des singuliers CVCVC ont un PI ‘aCCaaCun (cf. *Sanamun* / ‘*aSnaamun* ‘idole(s)’ – cf. Ratcliffe 1998 : 77). Il semble donc y avoir de bonnes raisons de considérer /a/ comme un morphème de pluralité, d’autant que cette voyelle est précisément celle qui apparaît au pluriel dans de nombreuses langues afroasiatiques (cf. Greenberg 1955).

L’obstacle auquel on se heurte aussitôt est que, comme le disent les pourcentages ci-dessus, /a/ n’est pas la seule voyelle : /u/ apparaît pour 40 % des singuliers CaCCun (cf. *qalbun* / *quluubun* ‘cœur(s)’, *najmun* / *nujuumun* ~ ‘*anjumun* ‘étoile(s)’), pour 25 % des CiCCun (cf. la variante *jusuumun* de *jismun* ‘corps’, et *Zillun* / *Zuluulun* ‘ombre(s)’, avec les variantes *Zilaalun* et ‘*aZlaalun*) et pour 18 % des CuCCun (cf. *quflun* / *qufuulun* ‘verrou(s)’). Certes, cette alternance /a/ ~ /u/ est en partie prédictible, car, semble-t-il, les PI en /u/ correspondent à des singuliers en /a/, donc à voyelle [–haute], de façon significativement plus fréquente qu’ils ne correspondent à des singuliers en /i/ ou /u/, donc à voyelle [+haute]. Un principe de polarité qualitative semble ainsi être à l’œuvre, tel que la voyelle du pluriel a tendance à inverser la valeur de hauteur de la voyelle du singulier (cf. Ratcliffe 1998 : 78).

Il y a toutefois une différence essentielle entre cette variation formelle et celle du yidin<sup>9</sup>. En yidin<sup>9</sup>, la forme du morphème est totalement prévisible à partir du mot lui-même ; en arabe, au contraire, déterminer la qualité de la voyelle du pluriel requiert que l’on connaisse la qualité de la voyelle correspondante du singulier, et cela sans aucune garantie de succès puisque, somme toute, 60 % des singuliers CaCCun, par exemple, *ne sont pas* associés à des pluriels en /u/.

Essaierons-nous, cela étant, de plutôt comparer l'arabe à des langues comme le latin ? Là encore, les conditions sont autres. Que le morphème pluriel ait pour forme au nominatif /-a/ plutôt que /-i/ ou /-ae/, c'est affaire de genre, neutre en l'occurrence. Mais le neutre est une propriété du lexème, quel qu'en soit le nombre, tandis que la valeur haute ou non haute de la voyelle en arabe est une propriété du seul singulier. Nous retombons ainsi sur le même problème de prévisibilité que précédemment. C'est là une autre façon de dire que les noms arabes ne sont pas répartis en classes morphologiques, comme ils le sont en latin, en russe, ou autre. Des formules comme CaCCun, CiCCun, etc., ne sont pas des classes, mais des descriptions auxquelles est attachée une certaine probabilité – le plus souvent inférieure à 50 % – que le pluriel aura telle forme. La diversité des PI, même s'il ne convient pas de l'exagérer, n'est donc pas la diversité ordonnée, ne s'opposant pas à une prévisibilité significative, que manifestent les langues à classes morphologiques.

L'identification de la prosodie iambique, ou non iambique, comme morphème-objet apparaît ainsi surtout comme une hypothèse destinée à sauver à tout prix cette notion dans ce cas particulier. Mais cette tentative de sauvetage est elle-même nécessaire vu le présupposé des auteurs que les PI sont formé *à partir* des singuliers leur correspondant, c'est-à-dire à partir d'une forme achevée et compacte dans laquelle il n'y a aucune place pour cet espace actif qu'est un site<sup>27</sup>. C'est ce présupposé qu'il convient de remettre en cause. Remise en cause qui, il faut le noter, ne contredit en rien l'idée défendue plus haut que les PI participent de la morphologie dérivationnelle, pour autant qu'il n'est nullement nécessaire que les dérivations aient pour bases des objets pleinement spécifiés. Or, les langues sémitiques nous mettent sous les yeux de façon particulièrement prégnante un objet réel partiellement spécifié, la racine. La thèse que je souhaite donc soutenir à présent est que les PI et les singuliers leur correspondant sont dérivés *parallèlement* à partir de la racine, et non pas ceux-là à partir de ceux-ci. Cette dérivation résulte, non de l'attachement d'un morphème-objet à une base, mais de l'activation d'un site dont la position est fixe et déterminée par rapport aux éléments de la racine, les consonnes radicales.

Cela étant, deux observations demeurent, qu'il nous faudra expliquer : (a) le gabarit minimal d'un 3C est CVCC, et il n'est jamais pluriel<sup>28</sup> ; (b) la forme finale d'un PI est quantitativement différente, par exemple, pour le nombre de mores, de celle du singulier correspondant. On a ainsi les rapports suivants : /kalb/ = 3 mores vs. /kilaab/ = 4 mores, /kitaab/ = 4 mores vs. /kutub/ = 3 mores, /jundub/ = 4 mores vs. /janaadib/ = 5 mores, /<sup>c</sup>ankabuut/ = 6 mores vs. /<sup>c</sup>anaakib/ = 5 mores, /<sup>i</sup>imbaraaTuur/ = 8 mores vs. /<sup>i</sup>abaaTirat/ = 6 mores (6 semble être le nombre maximal de mores d'un PI)<sup>29</sup>.

#### 4. Pour une théorie de la formation des PI comme activation d'un site

Trois hypothèses soutiennent cette théorie. La première concerne évidemment la présence d'un site dans la racine. Comme on l'a vu, cette

hypothèse prend sa source dans les travaux de Guerssel & Lowenstamm (1990). Ceux-ci localisent le site actif pour la flexion et la dérivation verbales de l'arabe entre la première et la deuxième consonne radicale ( $C_1$  et  $C_2$ ). À leur suite, Asfour (2001) place le site actif dans la formation des PI à droite de  $C_2$ , donc à gauche d'une  $C_3$  nécessairement présente dès lors que la racine est soumise à une opération morphologique (voir plus haut). Je reprends cette conclusion. On notera au passage que la théorie de McCarthy & Prince suppose implicitement l'existence d'un tel site, puisque la formation du pied iambique dans //kal/b/ > //kilaab/ implique l'insertion de matériel nouveau entre  $C_2$  et  $C_3$ . D'autre part, les études de Bohas (1997) montrent que  $C_3$ , quand elle est présente, n'appartient presque jamais à ce qu'il nomme « l'étymon », sorte d'archiracine commune à un ensemble de racines délimitant un champ sémantique<sup>30</sup>. Tout indique donc l'existence de deux domaines :  $C_1$ - $C_2$ , d'une part, le domaine de l'étymon, et  $C_3$ , d'autre part, et d'un espace entre eux, que je noterai ainsi :  $C_1C_2\{ \}C_3$ . C'est à l'intérieur de cet espace que se produisent les événements cruciaux pour la formation des PI.

Ma divergence d'avec Asfour porte sur le mode d'activation du site post- $C_2$ , comme je le nommerai désormais. Pour lui, il s'agit toujours de l'insertion d'une syllabe CV, qui a pour effet de créer une voyelle longue (voir plus loin). Cela lui permet de rendre compte des PI iambiques, mais le contraint à invoquer l'apophonie pour expliquer les PI qui ne manifestent pas cette prosodie. À la recherche d'une solution plus générale, je poserai au contraire que le site post- $C_2$  constitue un espace réellement libre, où l'activation morphologique peut faire apparaître une position C ou une position V. Dans le premier cas, une position V est automatiquement créée (voir ci-dessous).

Si aucune opération morphologique n'affecte la racine, le site post- $C_2$  demeure virtuel<sup>31</sup>. Dans ce cas, de par la contrainte CV (voir ci-dessous), la jonction entre  $C_2$  et  $C_3$  est occupée par un noyau plein ou vide. Il est évidemment vide si la racine est 2C (par exemple, *dam (un)* 'sang'), sauf à être identifié par la voyelle casuelle ; il l'est également dans *kalbun* et dans tous les noms du type CVCC. Il est plein en revanche dans *jabalun* 'montagne'. La situation qui nous intéresse est bien sûr celle où le site post- $C_2$  ne reste pas virtuel.

La deuxième hypothèse est due à Lowenstamm (1996), et elle affirme que CV est l'unique type syllabique. Une syllabe phonologique se compose donc nécessairement et exclusivement d'une attaque non branchante (C) et d'un noyau également non branchant (V)<sup>32</sup>. Ainsi, *kalb* (réalisation à la pause) a pour structure syllabique \$ka\$lv\$bv\$, où v note un noyau non spécifié (vide). C'est l'universalité de CV qui explique la création obligatoire d'une position V lorsque l'activation du site post- $C_2$  produit une position C. Lorsqu'elle produit une position V, en revanche, celle-ci n'a qu'à occuper la position V qui suit  $C_2$  de par le format syllabique pour qu'il en résulte une syllabe CV légitime, comme le montre le schéma : (a)...  $C_2v\{V_s\}$ ...  $\rightarrow$  (b) /... $C_2V_s$ .../, où (a) est une représentation morphologique, (b) une représentation phonologique,  $V_s$  la voyelle dans le site, et v un noyau vide donné par la contrainte CV.

La troisième hypothèse concerne la classe des segments appelés « glides », qui joue un rôle si important dans la morphologie de l'arabe. On considère traditionnellement qu'il existe deux semi-voyelles /y/ et /w/, correspondant aux voyelles hautes /i/ et /u/, alors même que ces segments appartiennent à la même classe de « glides » que l'occlusive et la fricative glottales /ʔ/ et /h/, généralement considérées comme des consonnes (cf. Ladefoged 1982; Dell 1985). Pour la théorie phonologique grossièrement esquissée au paragraphe précédent, il va de soi que les semi-voyelles ou glides ne sauraient former une classe indépendante : l'inventaire segmental contient les deux segments hauts /I/ et /U/ réalisés /i/ et /u/ en position de noyau syllabique, /y/ et /w/ en position d'attaque. Mon hypothèse est que cette même ambivalence caractérise le troisième sommet du triangle vocalique, à savoir le segment bas ([–haut]) que je noterai désormais /A/. Comme noyau, il se réalise /a/ (ou toute voyelle basse dans la zone de /a/ cardinal); comme attaque, il prend l'aspect d'une « gutturale » qui, selon les langues, se manifeste comme l'occlusive glottale /ʔ/ ou bien la fricative glottale /h/ (les deux glides bas), ou encore (mais pas en arabe) la fricative laryngale /ʕ/ (cf. Mazaudon 1997, Prunet 1996, 1998 pour des arguments empiriques et théoriques à ce propos). On aboutit ainsi à une asymétrie remarquable dans le système phonologique : tout segment qui peut se lier au noyau peut aussi se lier à l'attaque (étant donné qu'il en va également ainsi des segments moyens /E/ et /O/); mais l'inverse n'est pas vrai : il existe des segments qui se lient à l'attaque, mais qui ne peuvent se lier au noyau – et ce sont ceux-là précisément les vraies consonnes.

Cela étant, je propose l'hypothèse suivante : la pluralisation interne consiste en une amplification de la racine par insertion de substance phonologique liée à une position C ou V dans le site post-C<sub>2</sub>. « Amplification de la racine » : le terme entend signifier que la formation des PI n'est pas un processus affixal ; il ne s'agit pas d'un élargissement par un élément extérieur à la racine. Le processus est lexical pour autant qu'un couple singulier-PI correspond à deux ensembles d'éléments dans le lexique (voir plus loin), identiques si ce n'est pour la présence dans l'un de l'élément [pluriel]. C'est celui-ci qui provoque l'activation du site post-C<sub>2</sub>, faisant passer la racine à l'état amplifié. Les opérations morphologiques d'épel, dont nous verrons plus loin le détail, ne voient le site qu'activé par la présence de [pluriel]; elles y créent alors la substance phonologique qui matérialise l'amplification. La racine passe ainsi de l'état simple à l'état amplifié.

On dépasse ainsi les deux théories traditionnelles sur la formation des PI (cf. Petráček 1960), celle selon laquelle ils sont dérivés des singuliers correspondants, et celle selon laquelle PI et singuliers constituent deux items lexicaux autonomes. J'ai déjà montré l'inadéquation de la première, qui amène à supposer l'existence d'un morphème-objet peu plausible. Il est tout aussi évident que la seconde ne peut pas être strictement exacte, car il existe bien une relation spécifique entre les singuliers et les PI qui leur sont associés. Seulement, ce n'est pas la relation de type algébrique que l'on observe entre, par exemple, *dog* = x et *dogs* = x + y. C'est une relation fondée sur la *différence* entre les deux états de la racine, simple et amplifié, sans que le matériel phonologique réalisant

l'amplification soit lié *spécifiquement* au changement sémantique du singulier vers le pluriel, sans qu'il ait l'identité d'un morphème, en d'autres termes. Ce qui peut être considéré comme un morphème, en revanche, c'est le site lui-même et l'activité morphologique qui lui donne substance.

Revenons à présent à la propriété qu'a le site post-C<sub>2</sub> activé d'être occupé par une position C ou V. Cela signifie que la substance phonologique qui apparaît à l'emplacement du site doit être telle qu'elle puisse se lier à l'une ou l'autre position. Il doit donc s'agir de l'un des éléments /I/, /U/ ou /A/, qui ont seuls cette faculté. On explique ainsi que le site activé ne puisse en aucun cas se matérialiser sous la forme d'une obstruante, par exemple : celle-ci pourrait se lier à C, mais pas à V ; or, répétons-le, la position qui active le site est intrinsèquement C *ou* V, et ce qui s'y lie doit être apte à matérialiser l'un *ou* l'autre<sup>33</sup>.

De là proviennent aussi la diversité et la relative imprévisibilité des PI, car deux facteurs varient simultanément : la qualité de l'élément et la nature de la position, C ou V, dans laquelle il s'insère<sup>34</sup>. Cette variation est inhérente, en cela que, dans l'état présent de la langue, elle ne se laisse pas déduire de facteurs plus généraux. Avec les racines 3C, majoritaires, elle dépend presque entièrement de l'identité de la racine, et même alors le déterminisme n'est pas total, vu l'existence de variantes. Ce n'est qu'avec les racines 4C + que la formation est tout à fait prévisible. La présente hypothèse ne se propose donc pas de réduire, encore moins de bannir l'arbitraire des PI, car celui-ci est un fait, qui entraîne probablement des conséquences bien spécifiques pour l'acquisition. Ce qu'elle prétend faire, en revanche, c'est l'expliquer et montrer qu'il procède du fonctionnement d'un mécanisme unique, dont les parties sont déterminées même si leur combinaison ne l'est pas.

## 5. Application de l'hypothèse

Je veux montrer dans cette section que les processus fondamentaux qui viennent d'être décrits, plus quelques autres de moindre portée qui seront introduits à mesure que nous avancerons dans l'analyse, sont capables de produire un échantillon représentatif de PI et, potentiellement, toutes les formes existantes. Je commence par les items de racine 3C, puis nous passerons aux 2C et aux 4C +, pour finir par quelques formations particulièrement complexes. Au passage, une digression nécessaire montrera comment l'amplification de la racine touche un domaine plus large que les seuls PI, conformément aux propriétés qui, on l'a vu, distinguent le morphème-site du morphème-objet. Je ne chercherai pas dans cette section à déterminer ce qui distingue les items qui forment des PI de ceux qui n'en forment pas. Cela sera fait à la section suivante, dans le cadre d'un exposé détaillé concernant la forme de grammaire que suppose l'hypothèse ici illustrée.

Examinons tout d'abord la formation que McCarthy & Prince considèrent comme prototypique, à savoir ces PI iambiques dont *kalbun* / *kilaabun* est un bon exemple. Selon les hypothèses ci-dessus, la racine de cet item se note comme suit,



où les points représentent les positions V obligatoires de par le format CV, et les accolades enserrent le site post- $C_2$ , inactivé à ce stade :

(4)  $k.l.\{ \}b.$

Davantage qu'une notation de la racine, (4) est la représentation morphologique initiale de l'ensemble d'éléments qui constitue le lexème CHIEN<sup>35</sup>. Par «lexème», j'entends, conformément à l'usage, le commun dénominateur d'une collection de formes unies par une même dénotation, par exemple *dog* et *dogs* en anglais, *canis*, *canem*, *canibus*, etc. en latin (*cf.*, par exemple, Aronoff 1994). Je ne considère pas toutefois le lexème comme un paradigme, éventuellement unaire, de formes réelles, mais comme un ensemble d'éléments signifiants, abstraits au sens où ils n'ont pas de réalisation matérielle *dans le lexique* (voir section suivante). Au niveau lexical, (4) devrait donc être noté, par exemple,  $R_{klb}$ . Cette notation identifie les termes «lexème» et «racine» (R), comme il est permis de le faire dès lors que cela ne cause pas de confusion, et l'indice *klb* est une simple étiquette accolée à l'ensemble des éléments qui définissent le lexème-racine et le différencient de tous les autres. Par suite, (4) correspond à l'étape du composant morphologique où la racine vient à être interprétée, ce qui, en arabe, se traduit par l'identification de positions C, trois dans le cas présent<sup>36</sup>.

[Pluriel] est également un ensemble d'éléments, sans doute un ensemble singleton, non essentiellement différent de  $R_{klb}$ , si ce n'est qu'il ne lui correspond pas de dénotation, mais un concept, disons  $n > 2$ , où  $n$  est la cardinalité d'une collection d'exemplaires de n'importe quelle entité. On peut donc définir [Pluriel] comme fonctionnel, dans la mesure où il doit s'appliquer à des ensembles d'éléments dénotant des entités. L'application de [Pluriel], c'est-à-dire sa combinaison avec l'ensemble  $R_{klb}$ , selon des modalités qui seront précisées à la section suivante, est ce qui active le site post- $C_2$ , qui doit alors être occupé à l'étape correspondant à (4) par un membre de l'ensemble  $\{ /l/ \ /U/ \ /A/ \}$ , avec création d'une position C ou V. Dans notre cas, le membre est /A/, et nous allons supposer que la position créée est C. Celle-ci est nécessairement suivie d'une position V, comme expliqué précédemment, d'où :

(5)  $k.l.\{C_A.\}b.$  ( $C_A$  = position C réalisée /A/)

qui représente le résultat de l'amplification de la racine dans le cas du lexème  $R_{klb}$  combiné avec [Pluriel]. Il s'agit à présent d'identifier les positions V. Celle qui suit  $C_3$  /b/ l'est par la voyelle casuelle, éventuellement effacée plus tard. Comme on l'a vu, la position qui suit  $C_1$  /k/ est identifiée par une voyelle épenthétique de timbre /i/. Qu'en est-il des positions suivant  $C_2$  /l/ et / $C_A$ / amplificateur? L'hypothèse la plus simple, quasi nécessaire, est qu'elles sont identifiées par propagation du trait actif bas de /A/, soit :

(6)  $k.la\{C_Aa\}b.$



Il en résulte une suite consistant en un glide flanqué de deux voyelles identiques. Or, c'est une règle générale de la phonologie arabe que, dans un tel contexte, le glide s'efface. La suite /aC<sub>A</sub>a/ se résout ainsi en une voyelle longue, d'où la forme quasi finale /kilaab/. De fait, étant donné l'hypothèse CV, ce produit découle d'un principe universel : toute voyelle longue s'analyse en deux voyelles identiques séparées par le même élément en position C : /v<sub>i</sub>:/ = /v<sub>i</sub>C<sub>V<sub>i</sub></sub>v<sub>i</sub>/ (C<sub>V<sub>i</sub></sub> = position C réalisée V<sub>i</sub>)<sup>37</sup>.

Le glide amplificateur ne joue évidemment aucun rôle dans la vocalisation de la racine non amplifiée, c'est-à-dire au singulier. C'est l'identité du lexème qui en décide, de façon apparemment arbitraire. C'est pourquoi certains auteurs parlent de «signature» pour caractériser cette voyelle (cf. Levy 1971). Cela indique en tout cas – j'y reviendrai – que la seule vocalisation de la racine ne constitue pas une opération morphologique.

La même analyse s'applique à *najmun* / *nujuumun* 'étoile(s)', si ce n'est que le glide amplificateur y est /U/ au lieu de /A/, et que la voyelle épenthétique reçoit son timbre par harmonie vocalique (pour laquelle /A/ serait neutre)<sup>38</sup>. De même pour *jismun* / 'ajsaamun 'corps', moyennant le préfixe /'a/ (voir plus haut).

La comparaison de *jismun* / 'ajsaamun et de *kalbun* / *kilaabun* montre quelque chose d'important, à savoir que la qualité du glide amplificateur n'est pas liée à celle de la voyelle «signature» du singulier : toutes deux sont déterminées par l'identité du lexème, mais elles le sont de manière indépendante.

Soit à présent le cas de *kitaabun* / *kutubun* 'livre(s)'. Avant d'analyser la formation du PI, il convient de s'attarder sur quelques points d'égale importance que soulève la forme de cet item. Le singulier s'explique en effet par le fait que nous avons là un nom verbal du verbe de mesure III *kaataba* 'écrire à quelqu'un, correspondre avec quelqu'un', jamais employé, du reste, au sens de «fait d'écrire à quelqu'un, correspondance», ce pour quoi il existe un autre nom verbal, *mukaatabatun*, mais seulement comme nom simple voulant dire «livre» (anciennement «lettre, missive»). Voilà qui nous ramène à une observation déjà mise en valeur, à savoir la non spécialisation des gabarits pour ce qui est du nombre, le même gabarit CiCaaC étant associée au singulier dans *kitaabun*, mais au pluriel dans *kilaabun*.

Cela étant, la similitude formelle de /kitaab/ et /kilaab/ suggère une généralisation plus vaste, à savoir que le site post-C<sub>2</sub> est activé, non seulement par la présence de l'élément [pluriel], mais par d'autres éléments encore. Cette généralisation était du reste attendue : la nature même du morphème-site fait que l'activité morphologique qui lui donne substance est a priori susceptible d'une multitude de valeurs. Que ces dernières se limitent à la pluralité pour les noms est une donnée contingente. La morphologie verbale est plus riche.

Supposons que l'ensemble d'éléments associé à /kitaab/ inclue, outre R<sub>ktb</sub>, l'élément *v*, c'est-à-dire «verbalité», avec la valeur [III], ou encore [Applicatif], l'une des significations principales de la mesure III, plus l'élément *n*, c'est-à-dire «nominalité» (ces éléments, ainsi que la représentation ci-dessous, seront expliqués et précisés à la section suivante) :

(7)  $\{R_{ktb} \ v_{III} \ n\}$ 

En réalité, nous avons représenté là une nominalisation du type de *destruction* au sens processuel de « la destruction de Rome par les Vandales » (cf. Marantz 1997), si bien que (7) doit être associé à *mukaatabatun*, dérivation de niveau II qui forme un PE (*mukaatabaatun*). Nous savons pourtant que *kitaabun* s'emploie à la façon d'un nom primitif, ayant perdu toute valeur verbale. Autrement dit, nous avons là un exemple de ce processus de relexicalisation dont il a été question plus haut. On pourrait y voir l'effacement de l'élément *v*. Il est toutefois exclu que cet élément se perde complètement, car le fait est que les noms de gabarit CiCaaC ne s'associent qu'aux verbes de la mesure III. L'information portée par *v<sub>III</sub>* est donc indispensable.

Je supposerai donc que, sans disparaître, l'élément *v<sub>III</sub>* fusionne avec *n* (et nous verrons plus tard ce qu'il faut entendre précisément par là), si bien qu'il en résulte un élément complexe au sein duquel *n* est dominant pour le sens et la catégorie – puisque *kitaabun* ne garde rien du cadre sélectionnel de *kaataba*, à la différence de *mukaatabatun*. C'est dans cette inclusion de *v* dans *n* que consiste le processus de relexicalisation. On associera ainsi la représentation lexicale suivante à *kitaabun*<sup>39</sup>:

(8)  $\{R_{ktb} \ \{v_{III} \ n\}\}$ 

C'est donc la présence de l'élément complexe  $\{v_{III} \ n\}$  qui entraîne l'activation du site post-C<sub>2</sub>. Il est remarquable que le site soit alors occupé de la même manière que s'il était activé par [pluriel]. Il n'y a pourtant là rien d'étonnant, car, comme je l'ai montré plus haut, il n'existe pas d'autre manière de l'occuper.

Il n'est pas sans intérêt de noter à ce propos que l'on retrouve ainsi une observation ancienne, à savoir qu'il existe une ressemblance formelle entre certains types de PI et certains types de noms verbaux. Barth et d'autres auteurs de la fin du XIX<sup>e</sup> siècle et du début du XX<sup>e</sup> s'y sont fondés pour développer une théorie selon laquelle les PI (ou partie d'entre eux) étaient à l'origine des noms abstraits dénotant le concept général d'une activité ou d'un procès (ce que font en partie les noms verbaux), et où l'abstraction s'est muée en pluralité. J'ignore si cette théorie a valeur historique – la question est sans doute sans réponse – mais il est significatif qu'elle s'appuie sur un fait morphologique démontrable, à savoir que les PI et certains noms verbaux se forment par le même procédé (cf. Kihm 2003).

Tout aussi intéressant est le fait que cette analyse nous propose un corrélat substantiel de la distinction entre niveaux I et II, laquelle, on l'a vu, est en elle-même superflue. *Kitaabun*, du prétendu niveau I, met en jeu une fusion des éléments *n* et *v*, ce que ne fait pas *mukaatabatun*, du prétendu niveau II. Il reste à voir si cette corrélation est généralisable. J'y reviendrai.

Le processus décrit en (8) ne se limite pas aux noms verbaux. Soit, par exemple, le participe actif de mesure I *kaatibun*. Il peut avoir valeur verbale, verbo-nominale ou nominale. La représentation lexicale de la valeur verbale est celle-ci :

$$(9) \quad \{\{R_{\text{ktb}} n\} v_{\text{PA}}\}$$

L'élément  $n$  est ici subordonné à  $v$  vu la valeur verbale du participe, qui régit son complément à l'accusatif, non au génitif, alors même qu'il se pluralise (extérieurement) comme un nom et non comme un verbe, d'où l'inclusion nécessaire de  $n$  (cf. *hum kaatibuuna maktuuban* 'Ils sont en train d'écrire une lettre', à comparer à *hum yaktubuuna maktuuban* 'Ils écrivent une lettre').

Un participe actif de valeur verbo-nominale, d'autre part, régit son complément au génitif, formant avec lui un groupe nominal dit d'état construit, mais il ne peut former qu'un PE, et il conserve une signification processuelle qui justifie l'inclusion de  $v$  dans la représentation lexicale (cf. *hum kaatibuu l-makaatiibi* 'Ils sont ceux qui ont écrit les lettres', où *kaatibuu* est l'état construit de *kaatibuuna*). On peut donc en déduire que la hiérarchie des éléments est à l'inverse de (9),  $v$  étant subordonné à  $n$  :

$$(10) \quad \{\{R_{\text{ktb}} v_{\text{PA}}\} n\}$$

Enfin, *kaatibun* est un nom simple signifiant 'scribe'. Il forme alors le PI *katabatun*. On peut citer d'autres exemples du même type: *kaatibun* avec la signification 'écrivain' et le PI *kuttaabun*; *Taalibun* 'étudiant' (lit. «demandeur (de connaissance)») / *Talabatun* ~ *Tullaabun*; *faarisun* / *fawaarisu* 'cavalier(s)'<sup>40</sup>. Nous sommes là à nouveau dans le cas de *kitaabun*,  $n$  et  $v_{\text{PA}}$  ayant fusionné :

$$(11) \quad \{R_{\text{ktb}} \{v_{\text{PA}} n\}\}$$

Dans les trois représentations (9), (10) et (11) correspondant aux trois valeurs du participe, on constate que ce n'est pas la simple conjonction des éléments  $n$  et  $v_{\text{PA}}$  qui active le site post- $C_2$ , puisque la marque du participe actif de la mesure I est l'insertion de [a:], c'est-à-dire de /A/ dans le site «verbal» post- $C_1$ . Mais on constate également que la fusion de  $n$  et de  $v$  suffit pour que, [pluriel] étant ajouté, le site post- $C_2$  soit activé.

De cela, nous tirons la généralisation désirée plus haut quant à une définition substantielle des niveaux I et II du lexique nominal. Le niveau II correspond aux ensembles d'éléments où  $n$  et  $v$ , pourvus de diverses valeurs, et de quelque façon qu'ils se hiérarchisent, sont distincts. Le niveau I correspond aux ensembles d'éléments où  $n$  et  $v$  fusionnent (ou bien où seul  $n$  est présent, les noms primitifs). L'activation du site post- $C_2$ , quand elle n'est pas due à [pluriel], ne peut résulter que d'une fusion telle que  $\{v_{\text{III}} n\}$ . Lorsque les deux déclencheurs sont présents simultanément, par exemple au pluriel de *kitaabun*, on admettra que c'est l'élément le plus élevé dans la hiérarchie combinatoire (voir section suivante) qui agit (voir ci-dessous).

Pour conclure, encore un mot sur les noms verbaux, dont le comportement par rapport au pluriel présente une complexité qui mérite d'être signalée. Seuls les noms verbaux tout à fait séparés du verbe correspondant, tel *kitaabun*, forment un PI. Pour autant, il n'est pas vrai que les noms verbaux qui conservent une valeur verbale (par exemple, *mukaatabatun*) forment tous un PE. Il en est qui

ressemblent à des noms primitifs, par exemple *Darbun* 'coup' (cf. *Daraba* 'frapper'), mais dont la valeur verbale se voit au fait qu'ils s'emploient presque uniquement en fonction d'objets internes : cf. *Darabtuḥu Darban* 'Je lui ai frappé un coup', c'est-à-dire « Je l'ai frappé une fois ». Ceux-là ne se pluralisent pas directement, mais par le biais d'une forme de gabarit CaCaCatun, qui reçoit un PE féminin : cf. *Darabtuḥu talaata Darabaatin* 'Je l'ai frappé trois fois'. On peut en conclure que *Darbun*, comme *Darabatun*, est associé à une représentation lexicale où *n* et *v* sont distincts ; il devrait donc former un PE, ce à quoi sa morphologie de nom primitif s'oppose, d'où le « détour » par une forme plus adéquate<sup>41</sup>.

Revenons-en à présent au PI de *kitaabun*, *kutubun*. Nous venons de voir que le site post-C<sub>2</sub> est occupé au singulier. Mais nous avons également considéré l'hypothèse que [pluriel], en tant qu'élément plus élevé, prend le pas sur la fusion {*v<sub>III</sub>* n}. Cela signifie que, pour ce qui est de son activation par [pluriel], le site est vide. Il va donc être occupé par le glide que [pluriel] suscite, tout à fait indépendamment du singulier. Ce glide est à l'évidence /U/, et nous supposons que la position créée est V :

(12) k.t.{V<sub>U</sub>}b. (V<sub>U</sub> = position V réalisée /U/)

Les principes de syllabation garantissent que /U/ occupe le noyau suivant C<sub>2</sub> /t/ et se réalise [u] bref. Le noyau suivant C<sub>1</sub> /k/ est occupé par une voyelle épenthétique colorée [u] par harmonie vocalique, comme dans *njuumun* 'étoiles'. Un autre exemple est *mudunun*, PI du féminin marqué *madiinatun* 'ville'. Comme dans *kitaabun*, le site post-C<sub>2</sub> est occupé au singulier par un glide, ici /I/, en position C, d'où la voyelle longue, mais ce matériel est annulé par l'intervention de [pluriel]. En outre, l'exemple démontre l'invisibilité du genre du singulier pour la formation d'un PI. Autrement dit, le trait féminin du singulier, qui déclenche l'épel de la désinence féminine /- (a) t/, est lui aussi annulé par [pluriel] (voir section suivante).

Une variante de ce type est '*anjumun*, pluriel alternatif de *najmun* 'étoile'. Il se forme comme *kutubun*, si ce n'est pour le préfixe /'a/ dont la présence évite l'épenthèse. Une autre variante est *ḡurafun*, PI de *ḡurfatun* 'chambre'. Ici, le glide créant une position V dans le site est /A/. S'il est vrai que /a/ est neutre quant à l'harmonie vocalique, on attendait /i/ en épenthèse comme dans *kilaabun*. Ce semble toutefois être une caractéristique de cette sous-classe de PI que la voyelle suivant C<sub>1</sub> est identique à la voyelle correspondante du singulier<sup>42</sup>. Je reviendrai sur ce détail.

Nous examinons maintenant les lexèmes à racine 2C, par exemple *damun/ dimaa'un* 'sang(s)'. Une racine, on le sait, ne peut être traitée morphologiquement si elle ne comporte pas au moins trois positions C. Il s'ensuit que la racine « utile » de *damun* doit avoir la forme suivante :

(13) d.m.{ }.

où le point à la droite du site marque la C<sub>3</sub> virtuelle. Au singulier, le lexème n'étant soumis à aucun traitement morphologique particulier, la condition du triconsonantisme ne s'applique pas, si bien que la position n'est pas identifiée et n'a pas de réflexe phonologique. Au PI, le site post-C<sub>2</sub> est occupé par /A/ en position C, tandis que, la condition du triconsonantisme s'appliquant, C<sub>3</sub> se réalise par défaut comme /'/, l'occlusive glottale, d'où : <sup>43</sup>

(14) d.m.{C<sub>A</sub>.}'.

La propagation de /A/ aux positions V adjacentes et l'épenthèse de /i/ entre C<sub>1</sub> et C<sub>2</sub> produisent alors la forme finale *dimaa'un* 'sangs'.

Quant aux rares lexèmes à racine 1C que l'on range d'ordinaire avec les 2C parmi les racines anormales, on peut montrer que ce sont en fait des cas particuliers de racines 3C. Soit en effet *šaa'un* / *šiyaahun* ~ *šiwaahun* 'mouton(s)'. Le [a:] long et l'occlusive glottale trahissent la présence de /A/ en C<sub>2</sub> et C<sub>3</sub>, d'où la représentation ci-dessous de la racine <sup>44</sup>:

(15) š.C<sub>A</sub>.{ }C<sub>A</sub>.

La raison habituellement avancée pour soutenir que la racine de *šaa'un* est 1C est que seul /š/ se retrouve au singulier et au pluriel. Avec une représentation comme (15), cependant, on voit que la racine peut très bien être analysée comme 3C, C<sub>2</sub> et C<sub>3</sub> étant des glides. Le site post-C<sub>2</sub> est vide en (15). Il est activé au pluriel et rempli par /A/ en position C. Soit maintenant une généralisation de ce que nous avons établi pour *dimaa'un* et que les grammairiens traditionnels ont reconnu depuis longtemps, à savoir que les glides radicaux sont des segments faibles, des occupants par défaut de positions consonantiques, qui, par cela même, ne sont pas contraints de se réaliser identiquement, voire de se retrouver dans ces formations parallèles que sont le singulier et le PI. C'est ce qui est illustré ici : au PI, le glide en C<sub>2</sub> est /I/ ou /U/, en variation libre, au lieu de /A/, qui se réalisent /y/ ou /w/ puisqu'ils ne se trouvent pas entre voyelles identiques. Le glide en C<sub>3</sub>, quant à lui, garde la valeur /A/, mais il se réalise /h/, c'est-à-dire la fricative au lieu de l'occlusive glottale, valeur *a priori* possible, mais rare, peut-être limitée à ce contexte. D'où :

(16) š.C<sub>I/U</sub>.{C<sub>A</sub>}C<sub>A=h</sub>.

réalisé *šiyaahun* ou *šiwaahun* (/i/ épenthétique).

Les PI des noms à racine 4C + ne présentent pas de difficultés particulières. Prenons *jundubun* / *janaadibu* 'sauterelle(s)', l'exemple de McCarthy & Prince. Il y a une chose intéressante à noter, déjà mentionnée au passage, à propos de la vocalisation du singulier : seuls les 4C primitifs (non dérivés) présentent cette suite de deux /u/ au singulier ; dans les 3C, /u... u/ ne se rencontre qu'au pluriel (cf. *kutubun*, *mudunun*, etc.). Il se peut que cela soit en relation avec le fait que les 4C primitifs sont en majorité, voire tous, des emprunts à des langues non sémitiques (cf. Petráček 1960). S'il en est bien ainsi, cela

voudrait dire que, dans les noms autochtones, /u/ en  $V_2$  précédé de /u/ en  $V_1$  s'interprète toujours comme la réalisation de /U/ amplificateur dans le site post- $C_2$ . Voilà qui ferait de ce /u/ un signal de pluralité, mais, pour les raisons exposées plus haut, ne suffirait pas à en faire un morphème (-objet) de pluralité, du moins synchroniquement. L'unicité du gabarit des PI pourrait également être liée à leur origine étrangère.

La représentation morphologique de la racine de *jundubun* est la suivante :

(17) j.n.{ }d.b.

Il y a deux consonnes à droite du site post- $C_2$ , si bien que la voyelle qui les sépare peut être /u/ sans risque d'être confondue avec la réalisation de /U/ amplificateur, puisqu'elle est extérieure au site. Le PI se forme simplement en occupant le site post- $C_2$  par /A/ en C :

(18) j.n.{ $C_A$ .}d.b.

Le /i/ entre  $C_3$  /d/ et  $C_4$  /b/ est clairement épenthétique. Je n'ai pas d'explication, en revanche, pour le /a/ entre  $C_1$  /j/ et  $C_2$  /n/ (pourquoi pas \*/jinaadib/ ?). Il se peut que ce /a/ soit la copie de /A/ – mais pourquoi copie plutôt qu'épenthèse ?

Quant à la chute de la cinquième consonne dans les noms à racine 5C, nous avons vu qu'elle est due au fait que la taille maximale de la racine utile (morphologiquement traitable) est 4C. On notera toutefois que l'élagage d'une « vraie » consonne, c'est-à-dire d'une obstruante, n'intervient, semble-t-il, qu'en dernier recours : on préférera toujours, si c'est possible, supprimer un glide, voire une sonante, même s'il ne s'agit pas de la dernière consonne de la racine (voir l'exemple à la note 18). J'y reviendrai.

Nous allons maintenant conclure cette section par l'examen de deux cas un peu plus complexes. Le premier est *diibaajun* / *dayaabijju* ~ *dabaabijju* 'brocart(s)'. Bien qu'il semble que nous ayons affaire à une racine 3C au singulier (*dbj*), le pluriel montre une racine 4C, avec le PI associé. Cette racine a deux variantes, l'une avec un glide en  $C_2$  (*dybj*), l'autre dans laquelle  $C_2$  est identique à  $C_3$  (*dbbj*). En fait, vu nos hypothèses sur la nature des voyelles longues, la racine du singulier *diibaajun* doit être celle-ci :

(19) d.  $C_1$ .{ }b.  $C_A$ .j.

soit une racine 5C dont deux des radicales,  $C_2$  et  $C_4$ , sont des glides. Nous savons qu'une radicale doit disparaître au PI, et que, dans le cas présent, ce sera l'un des deux glides plutôt que la  $C_5$  obstruante /j/ (voir ci-dessus). Supposons à présent que, lorsque un glide est élagué au PI afin de satisfaire la condition régissant la taille maximale de la racine utile, la position qu'il occupe peut être préservée<sup>45</sup>. Cette hypothèse, associée au choix que nous avons d'élaguer l'un ou l'autre glide, va nous permettre d'expliquer l'alternance des deux PI. Dans l'un et l'autre, le site post- $C_2$  est occupé par /A/ en C.

Soit le cas où le glide élagué est /I/ ; il en résulte la forme suivante :

(20) d.C.{C<sub>A</sub>.}b.C<sub>A</sub>j.

où C marque la position évacuée par /I/. Cette position doit être identifiée: elle l'est par copie des traits de C<sub>3</sub> /b/, d'où /d.b.{C<sub>A</sub>.}b.C<sub>A</sub>j/. Comment expliquer le /i:/ long à la place de /A/ radical? La solution retenue plus haut pour *šaa'un* s'applique ici, avec une légère variante: le glide, segment faible, peut avoir des identités différentes au singulier et au PI; le PI des 4C + met en jeu une voyelle épenthétique /i/, c'est-à-dire la réalisation vocalique de /I/, entre C<sub>3</sub> et C<sub>4</sub> (cf. *janaadibu*); ici, /I/ s'ancre à la position occupée par C<sub>4</sub> /A/, que la faiblesse du glide lui rend accessible, produisant automatiquement une voyelle longue /i:/. On notera que, de ce fait, la forme apparente, *dabaabiiju*, redevient quadriconsonantique. J'y reviendrai.

Voyons à présent l'autre possibilité: le glide élagué est /A/. Il s'ensuit la forme suivante au PI:

(21) d.C<sub>1</sub>{C<sub>A</sub>.}b.C.j.

Après insertion de /a/ entre C<sub>1</sub> et C<sub>2</sub> et propagation du trait bas de /A/ amplificateur aux positions adjacentes, /I/ radical se trouve bien entre des voyelles identiques, mais son effacement produirait une voyelle surlongue /aaa/. Il est donc réalisé /y/, puisque en position d'attaque syllabique. Le /i:/ long entre /b/ et /j/, reconvertis en C<sub>3</sub> et C<sub>4</sub>, a la même origine que dans la forme concurrente. La réalisation de (21) est donc *dayaabiiju*.

Le deuxième cas est celui de *taa'jirun* / *tujjaarun* 'commerçant(s)'. Il s'agit d'un participe actif relexicalisé (cf. *kaatibun* / *kuttaabun* 'écrivain(s)'), dont la représentation morphologique initiale (après relexicalisation) est donc:

(22) t.(C<sub>A</sub>.)j.{ }r.

Les parenthèses notent que /A/, en tant que formant du participe, ne compte pas comme consonne radicale, d'où la position du site et le fait que *taa'jirun* se pluralise comme un 3C, à la différence de *diibaajun*. Le problème à résoudre est la gémination de C<sub>2</sub> dans la forme de PI. Les hypothèses utilisées dans le cas précédent donnent la solution: au PI, où le site est occupé par /A/ en C, la position C du /A/ participial est conservée, comme l'est celle de /I/ ou de /A/ dans *diibaajun*, d'où:

(23) t.C.j.{C<sub>A</sub>.}r

Cette position est alors identifiée par /j/, analysable comme C<sub>2</sub> ou comme C<sub>3</sub>, à nouveau comme dans *dabaabiju*, d'où /t.j.j.{C<sub>A</sub>.}r/. On peut donc généraliser: la position est identifiée par la consonne radicale la plus proche à sa droite, c'est-à-dire par la consonne en position de la gouverner phonologiquement (cf. Kaye, Lowenstamm & Vergnaud 1990). La position vocalique brève entre consonnes identiques s'efface en vertu d'une règle phonologique générale, d'où *tujjaarun* avec gémination de C<sub>2</sub>, qui rend la racine pratiquement 3C. La seule chose qui

demeure sans explication, mais pas seulement dans ce cas, est la valeur /u/ de la voyelle épenthétique.

On notera pour finir que l'identification de la position consonantique correspondant à un glide au singulier par la consonne gouvernante n'est pas la seule solution : cette position peut aussi être réalisée directement comme une occlusive glottale, c'est-à-dire la consonne par défaut. On explique ainsi des PI tels que *Haqaa'ibun* 'valises', associé à *Haqaybatun*, ou *sajaa'irun* 'cigarettes', associé à *siijaaratun* (je laisse au lecteur le soin d'effectuer les opérations). L'existence de ces formes est une confirmation supplémentaire de la présente analyse.

## 6. Conséquences pour l'architecture de la grammaire

La théorie de la formation des PI qui vient d'être illustrée peut être dite *morpholexicale*. Elle affirme en effet que cette formation résulte de l'association de deux opérations : d'une part, la combinaison intralexicale de l'élément [pluriel] avec un ensemble d'éléments définissant un lexème-racine ; d'autre part l'amplification de cette racine épelée en morphologie par l'insertion d'un glide dans le site post-C<sub>2</sub>. La première opération suppose que le lexique est, non pas un répertoire de formes, mais une collection d'éléments signifiants et fonctionnels combinables en ensembles finis, correspondant aux concepts que le langage exprime. Ces ensembles d'éléments – terme qui rend le *feature bundle* de la DM dont je reprends en grande partie les hypothèses (cf. Halle & Marantz 1993 ; Marantz 1997) – sont associés à des représentations phonologiques dans la morphologie, se muant ainsi en vocables (*Vocabulary items*)<sup>46</sup>.

Idéalement, cela suffirait. Il existe toutefois des éléments dont l'apparition dépend entièrement de la distribution syntaxique de ces associations ensembles – vocables, par exemple le cas structural, nominatif ou accusatif, du sujet ou de l'objet direct. C'est donc que la syntaxe intervient (ou peut intervenir) dans la connexion entre le lexique et la morphologie. Je n'entre pas dans ce débat complexe, qui touche entre autres choses au fait de savoir si la morphologie doit être conçue comme scindée (*split*) ou non, et si et dans quelle mesure les catégories fonctionnelles de cas, de nombre, etc., sont représentées comme autant de têtes syntaxiques (pour une revue détaillée, quoique un peu ancienne, des questions, cf. Spencer 1991). Il me suffira de faire valoir que mon analyse de la formation des PI ne donne aucun rôle à la syntaxe. Elle évite en particulier tout recours à une projection syntaxique du nombre (NumP), à la tête Num° de laquelle la tête nominale N° s'adjoindrait par déplacement<sup>47</sup>. Comme on le verra, les opérations intralexicales de combinaison des éléments accomplissent tout ce que la transformation de déplacement des têtes est censée produire, du moins dans ce cas, et cela d'une manière théoriquement plus adéquate étant donné la nature dérivationnelle de la pluralisation évoquée plus haut (cf. Beard 1995).

S'il est vrai, pour les raisons mentionnées à l'instant, que la syntaxe constitue un passage obligé entre le lexique et la morphologie, on doit donc dire



que l'insertion syntaxique (comme tête d'une projection NP) d'un ensemble d'éléments épelé comme un PI n'a aucun effet sur la constitution *interne* de cet ensemble (c'est-à-dire abstraction faite de l'élément de cas, par exemple, qui s'y ajoute du fait de l'insertion)<sup>48</sup>.

Cette idée est importante, car elle fonde la possibilité de la théorie des PI défendue ici. Pour la DM «classique», tout ensemble d'éléments (*feature bundle*) lexicaux doit être associé à un morphème-objet distinctif<sup>49</sup>. Cette association nécessaire résulte de l'hypothèse que les ensembles d'éléments fonctionnels, par exemple [Accusatif, Pluriel], sont manipulés par la syntaxe à l'instar des ensembles d'éléments radicaux, et livrés à la morphologie à la sortie (*output*) des opérations syntaxiques. Celles-ci étant par nature concaténatives, procédant par adjonction de têtes, il s'ensuit qu'un ensemble d'éléments morphologiquement visible doit correspondre à une suite délimitée de segments phonologiques (par exemple, en simplifiant beaucoup, [Accusatif, Pluriel, Déclinaison II] = /os/ en latin). L'avantage de ce dispositif est qu'il produit la position respective des morphèmes sans qu'il soit besoin de la spécifier: /os/ est un suffixe parce que *pueros* est formé par déplacement de la tête radicale épelée /puer/ pour l'adjoindre, à gauche par définition, à la tête fonctionnelle contenant l'ensemble d'éléments épelé /os/<sup>50</sup>. (En fait, le modèle prédit, sans doute à juste titre par ailleurs, que tout morphème flexionnel doit être un suffixe, d'où la nécessité d'opérations spéciales telles que la fission – cf. Halle (2000) – pour produire les flexions préfixales.)

Son inconvénient est qu'il traite difficilement les modifications morphologiques, flexionnelles ou non, auxquelles ne correspond pas une suite délimitée de segments, par exemple les pluriels apophoniques de l'anglais tels que *foot* /*feet*, *mouse* / *mice*, etc., où il serait absurde d'analyser les segments /i:/ et /a/ comme des morphèmes. La solution habituellement retenue (cf. Halle & Marantz 1993) consiste à supposer la présence d'un suffixe zéro ayant la propriété de déclencher l'apophonie. Pour les pluriels anglais, cette hypothèse a le mérite de l'exactitude historique, puisqu'elle revient à retrouver l'umlaut originel dû à une voyelle haute finale sous ce que la disparition de celle-ci a mué en apophonie (*Ablaut*). On ne voit guère son intérêt pour les PI arabes, les processus internes mis en jeu n'étant pas de ceux qu'un élément extérieur a jamais pu déclencher au niveau phonologique<sup>51</sup>.

Concevoir la pluralisation interne comme une correspondance entre opérations lexicales et morphologiques, en revanche, supprime le recours nécessaire à des morphèmes-objets (les *morphemes as things* de Spencer 1991); elle peut être l'expression de processus morphologiques (*morphemes as processes* – cf. Anderson 1992) formalisables par des fonctions (cf. Raffelsiefen 1992) et dont le morphème-site, association d'un lieu et d'une activité, est un cas particulier. Afin de le montrer, il va nous falloir (a) expliciter la notion d'opération intralexicale; (b) formaliser une théorie de l'épel morphologique qui fasse autre chose que concaténer des morphèmes-objets. C'est à quoi je vais consacrer les deux sous-sections ci-dessous, en me cantonnant aux éléments nécessaires et

suffisants pour rendre compte de la formation des PI telle qu'elle a été décrite à la section précédente, et pour expliquer en quoi ils se distinguent des PE.

### 6.1. Les opérations intralexicales

Le lexique, on l'a vu, est ici conçu comme une collection d'éléments qui se combinent en ensembles finis auxquels on peut donner le nom de lexèmes. Parmi ces éléments, il y a ceux dont la combinaison produit une racine, c'est-à-dire un objet lexical auquel est associée une dénotation (le lexème minimal). Pour une langue comme l'arabe, on notera les racines par un R étiqueté d'une suite de consonnes à interpréter comme un pur indice, par exemple  $R_{klb}$  correspondant au concept de «chien»,  $R_{ktb}$  correspondant à un concept moins concret exprimable comme «lié au fait d'écrire», etc.<sup>52</sup>. Comme le montre ce dernier exemple, les racines sont dépourvues de catégorie syntaxique<sup>53</sup>:  $R_{ktb}$  est sous-jacente à *kataba* 'il a écrit' comme à *kitaabun* 'livre', tout comme  $R_{destruc}$  est sous-jacente à *détruire* et à *destruction*<sup>54</sup>. Même une racine comme  $R_{klb}$ , qu'on pourrait croire exclusivement nominale, soutient non seulement le nom *kalbun* 'chien', mais aussi le verbe *kaliba* 'être ou devenir enragé'.

Il s'ensuit que tout nom est la nominalisation, et tout verbe la verbalisation d'une racine. On est ainsi amené à supposer l'existence de deux éléments, *n* et *v*, qui valent respectivement pour «nominalité» et «verbalité»<sup>55</sup>. Je suis donc Marantz (1997) sur ce point, mais je m'écarte de lui lorsque je considère que *v* et *n*, en tant que verbalisateur et que nominalisateur, n'ont pas à se projeter dans la syntaxe, mais se combinent lexicalement aux racines. Il me semble en effet que Marantz ne tire pas toutes les conséquences de cet apport essentiel de la DM qu'est la réanalyse du lexique en une composante *sémantique*, en fait la composante sémantique du dispositif grammatical. Formalisées comme il convient (voir plus loin), les opérations lexicales n'ont ainsi plus rien de «lexical» au sens habituel, c'est-à-dire d'idiosyncrasique, de «privé» pour paraphraser le titre de l'article de Marantz<sup>56</sup>. Cette réinterprétation de *v* et de *n* comme des éléments lexicaux permet en outre une réduction significative du nombre et, surtout, des types des catégories fonctionnelles présentes en syntaxe. La limitation des types est le plus important, car il est n'est évidemment pas question, comme je l'ai déjà laissé entendre, d'exclure que les éléments lexicaux puisse se combiner en ensembles au terme d'opérations syntaxiques. Le cas structural a été cité, il se peut que la combinaison du temps avec une base verbale entre aussi dans ce que la syntaxe doit accomplir<sup>57</sup>. Je ne vois en revanche aucune raison pour que la syntaxe ait toujours à prendre en charge la combinaison du nombre avec une base nominale. Les faits d'accord et d'autres – cf. l'analyse des génitifs en gallois et en hébreu de Nash & Rouveret (1997) – ne constituent pas un argument en ce sens, car ils n'impliquent rien de plus que l'accessibilité du trait de nombre depuis l'extérieur du lexème source, qui peut être assurée de quelque façon que le nombre et la base nominale se combinent.

Vu leur manifestation en arabe (et en français), j'admettrai que  $n$  et  $v$  sont des éléments fonctionnels – mais il y a des raisons de penser qu'ils peuvent aussi être radicaux (cf. Kihm à paraître (a)). Comme toute unité,  $n$  et  $v$  peuvent être notés comme des ensembles unaires, soit  $\{n\}$  et  $\{v\}$ . Cette conversion permet de généraliser l'hypothèse selon laquelle le lexique est fait d'ensembles d'éléments : le lexique est un ensemble d'ensembles, et il ne contient que des ensembles.

Dans les faits,  $\{v\}$  n'est sans doute jamais un ensemble singleton, s'il est vrai qu'il détermine non seulement la verbalité du lexème fondé sur la racine, mais aussi, par exemple, le type de verbe (transitif, inaccusatif, etc.)<sup>58</sup>. On a donc un ensemble multiple  $\{v_i, v_j \dots v_n\}$  dont les éléments sont des valeurs de  $v$ . Il en va de même de  $\{n\}$  dans les langues qui *classifient* les noms. Les valeurs de  $n$  dans l'ensemble  $\{n_i, n_j \dots n_n\}$  sont alors des classes : genres, classes nominales, etc. (cf. Kihm à paraître (a) ; pour les procédés classificatoires en général, cf. Aikhenvald 2000)<sup>59</sup>. En arabe, il existe deux genres traditionnellement dénommés « masculin » et « féminin ». L'ensemble  $\{n\}$  est donc binaire :  $\{n, n_f\}$ . Tout ensemble lexical comprenant  $n_f$  correspond à un nom féminin, tout ensemble comprenant  $n$  non valué correspond à un nom masculin (le genre non marqué – pour l'origine du genre en arabe et en sémitique, cf. Corriente 1971 : 55 et suiv.).

Vu le contenu de  $\{n\}$  et  $\{v\}$ , la première combinaison intralexicale doit être celle d'une racine, elle-même un ensemble, avec l'un ou l'autre ensemble<sup>60</sup>. De ce caractère ensembliste des éléments lexicaux, il suit que leurs combinaisons doivent être formalisables comme des fonctions appliquées aux ensembles (cf. Cooper 1974 ; Partee *et al.* 1990). Deux fonctions paraissent constituer le minimum requis, et il ne semble pas qu'il y en ait besoin de davantage. L'une, l'*addition*, consiste à former un nouvel ensemble à partir de deux ensembles donnés :

$$(24) \quad \{A\} \{B\}_{ad} \rightarrow \{\{A\}, \{\{A\}, \{B\}\}\} = < \{A\}, \{B\} >$$

Cette fonction prend deux ensembles  $\{A\}$  et  $\{B\}$  à  $n$  éléments ( $1 \leq n$ ) et produit un ensemble formé de  $\{A\}$  et de l'ensemble incluant  $\{A\}$  et  $\{B\}$ , lequel est égal à une paire ordonnée  $< x, y >$  telle que tous les éléments de  $\{A\}$  précèdent tous les éléments de  $\{B\}$ . Étant donné un isomorphisme entre les opérations du lexique et les opérations de la morphologie, cette fonction nous permet de rendre compte des cas où l'exposant d'un élément fonctionnel apparaît suffixé à l'exposant d'un élément radical. Soit  $\{A\}$  une racine  $R_x$ , et soit  $\{B\}$  l'élément  $\{n_f\}$  de  $\{n\}$  : (24) est ainsi associée aux féminins marqués tels que *madiinatun* /madiin-at-f-u<sub>k</sub>-n<sub>idf</sub>/ 'ville' ( $k$  = cas, *idf* = indéfini).

La fonction addition suppose crucialement une *hiérarchie* des activations lexicales : en (24), l'activation de la racine est surordonnée par rapport à celle de l'élément fonctionnel. Autrement dit, la racine est le *domaine* de la fonction. On souhaite évidemment faire une généralité de cette proposition, et dire qu'il en est toujours ainsi, sans doute parce que la racine est l'élément porteur de dénotation<sup>61</sup>. De cette façon, la fonction addition nous donne tout ce que la transformation « déplacement de tête » est censée assurer, y compris la nature par principe

suffixale des morphèmes non radicaux distincts (voir plus haut et note précédente). On notera au passage que la fonction ne fait pas référence à la distinction entre flexion et dérivation<sup>62</sup>. J'y reviendrai.

La deuxième fonction dont nous avons besoin est l'*inclusion*, qui consiste à inclure un ensemble donné dans un autre ensemble également donné :

$$(25) \quad \{A\} \{B\}_{incl} \rightarrow \{A \{B\}\}$$

Contrairement à l'addition, cette fonction ne produit pas une paire ordonnée. Il s'ensuit que l'exposant potentiellement associé à  $\{B\}$ , l'élément fonctionnel, ne peut être linéarisé par rapport à l'exposant associé à  $\{A\}$ , la racine, ce qui entraîne que  $\{B\}$  ne peut en fait être exprimé par un morphème-objet, car ceux-ci doivent être linéarisés. (Cette proposition n'est autre que le *Linearization Correspondence Axiom* ou LCA de Kayne 1994<sup>63</sup>.) On rend compte ainsi, par inclusion de  $\{n_f\}$  dans  $R_x$ , des féminins « par convention » tels que *šamsun* /šams<sub>F</sub>-u<sub>k</sub>-n<sub>idf</sub>/ 'soleil'<sup>64</sup>.

J'ai dit que la combinaison, par addition ou inclusion, de  $n$  (ou de  $v$ ) avec une racine devait être la première opération. Je vais renforcer cette hypothèse en faisant la proposition suivante pour les lexèmes nominaux (j'ignore si elle est généralisable aux lexèmes verbaux) : tout autre élément  $e$  lexicalement combiné avec une racine se combine avec  $n$ , et c'est la combinaison  $\{n, e_x\}$  qui se combine avec la racine. L'avantage théorique est de garantir la binarité des applications fonctionnelles. Les avantages descriptifs apparaîtront, je l'espère, au cours de l'exposé.

L'autre élément en question est la pluralité, et il est le seul. En effet, le cas en arabe est uniquement structural, si bien que son épel est une fonction de la correspondance entre le lexème comme tête syntaxique et la morphologie (voir plus loin)<sup>65</sup>. Quant au tanwîn, en tant qu'il signifie l'indéfinitude, il marque une propriété du groupe nominal (DP) dont le lexème est la tête. Cela étant, il faudrait évidemment rendre comptes des syncrétismes qui affectent ces éléments (par exemple, au PE masculin, nominatif /-uuna/ vs oblique /-iina/).

Jusqu'à présent, nous avons implicitement considéré le pluriel comme un trait affectant un lexème, de dénotation singulière en son absence. Cette conception ne s'accorde pas avec le système d'éléments désormais adopté. On supposera plutôt l'existence d'un concept de nombre lexicalement représenté par un ensemble dont les éléments sont des valeurs de cardinalité. Pour l'arabe, trois valeurs sont nécessaires :  $\{1, 2, 3 +\}$ , correspondant aux catégories traditionnelles du singulier, du duel et du pluriel. On admettra que, sans doute de façon universelle,  $\{1\}$  s'inclut dans  $n$ , d'où  $\{n \{1\}\} = n \times 1 = n$ . On peut donc ne pas noter cet élément<sup>66</sup>. Le duel  $\{2\}$  s'additionne à  $\{n\}$ , et l'ensemble  $\{\{n\}\{2\}\}$  s'additionne à la racine, d'où par exemple,  $\{\{R_{mlk}\}\{n\}\{2\}\} = malikaani$  'deux rois',  $\{\{R_{mlk}\}\{n_f\}\{2\}\} = malikataani$  'deux reines'.

On a déjà compris que les PI expriment l'inclusion de l'inclusion  $\{n \{3 +\}\}$  dans la racine. Il convient toutefois d'en dire plus, car nous devons rendre compte des faits d'accord. Je supposerai donc que l'inclusion de  $\{3 +\}$  – désormais noté  $\{pl\}$  pour simplifier – dans  $\{n\}$  est totale, en ce sens que  $\{pl\}$  inclus dans  $\{n\}$  devient la valeur de  $\{n\}$  :  $\{n_{pl}\}$ . Cela peut s'exprimer autrement,

si l'on admet que la notation informelle  $\{n_f\}$  vaut pour  $\{n \{f\}\}$ , et  $\{n\}$  pour  $\{n \{ \} \}$ ,  $\{ \}$  l'ensemble vide interprété comme le genre non marqué. On dira alors que  $\{pl\}$  se substitue à  $\{f\}$  ou à  $\{ \}$ . Il s'ensuit qu'un lexème incluant  $\{n_{pl}\}$  ( $= \{n \{pl\}\}$ ) n'est pas valué pour le genre. Mais il s'ensuit aussi, du moins peut-on en faire l'hypothèse, que  $\{pl\}$  inclus dans  $\{n\}$ , lui-même inclus dans la racine, n'est pas accessible aux processus d'accord. On explique ainsi que l'accord, par exemple d'un adjectif, avec un PI se fasse au féminin singulier (cf. *mudunun kabiiratun* 'des grandes villes') : au singulier par défaut, et au féminin parce que celui-ci a aussi valeur de neutre<sup>67</sup>. La représentation lexicale générique d'un PI est donc :

(26)  $\{R_x \{n_{pl}\}\}$  (*kilaabun, mudunun, etc.*)

Les PE «féminins» partagent avec les PI la propriété d'entraîner un accord au féminin singulier<sup>68</sup>. Les guillemets se justifient par le fait que la désinence /-aat/ ne s'attache pas, tant s'en faut, qu'aux lexèmes féminins (voir section 2). On peut en conclure que ces PE manifestent le même processus d'inclusion de  $\{pl\}$  dans  $\{n\}$  que les PI, si ce n'est que l'ensemble résultant  $\{n_{pl}\}$  est additionné, et non pas inclus, à la racine. D'où la représentation générique suivante :

(27)  $\{\{R_x\}\{n_{pl}\}\}$  (*malikaatun, ijtimaa'aatun, etc.*)

Restent les PE masculins, qui entraînent l'accord «direct», soit que l'adjectif prenne la forme du PE masculin (par exemple, *mu'allimuuna gaa'ibuuna* 'des professeurs absents'), soit qu'il forme un PI (par exemple, *mu'allimuuna kibaarun* 'des professeurs âgés').  $\{pl\}$  est donc accessible pour l'accord, ce qui signifie qu'il se combine avec  $\{n\}$  par addition, non par inclusion, et que l'ensemble résultant s'additionne à la racine :

(28)  $\{R_x\}\{\{n\}\{pl\}\}$  (*mu'allimuuna*)

On voit que la différence des PI et des PE, telle que manifestée par leur constitution morphologique et les particularités de l'accord qu'ils commandent, peut être entièrement caractérisée par le jeu des deux fonctions lexicales d'addition et d'inclusion appliquées à des éléments réunis en ensembles. Ce que nous n'avons pas encore expliqué, en revanche, c'est pourquoi telles racines requièrent l'application de l'une ou de l'autre fonction. En d'autres termes, qu'est-ce qui, dans ce système, correspond à la division du lexique en niveau I et niveau II ?

Reprenons nos remarques sur cette division et reformulons-les au moyen des concepts que nous venons de développer; nous arrivons à la proposition suivante : l'inclusion de  $\{n_{pl}\}$  dans un ensemble lexical n'est possible que si celui-ci est disjoint de tout paradigme (originellement ou par relexicalisation). D'autre part, nous avons constaté que  $\{n_{pl}\}$  peut s'inclure (PI) ou s'additionner (PE «féminins»), mais que  $\{\{n\}\{pl\}\}$  ne peut que s'additionner (PE masculins). La conjonction de ces deux observations nous donne la réponse : un seul élément peut être inclus dans une racine. Appelons cela le principe de l'inclusion unique (PIU) :

(29) PIU :  $*\{\{R_x \{e_y\}\}\{e_z\}\}$  sauf si  $\{e_y\} \{e_z\}_{incl} \rightarrow \{e_y \{e_z\}\}$

Autrement dit : on ne peut inclure un élément  $e_z$  dans une racine  $R_x$  où est déjà inclus un élément  $e_y$ , sauf si  $e_z$  est inclus dans  $e_y$ , l'ensemble résultant valant alors pour un seul élément, à l'instar de  $\{n_{pl}\}$ .

Soit le participe actif de mesure II *mu'allimun* 'enseignant, professeur'. Sa formule lexicale met en jeu l'inclusion dans la racine  $R_{clm}$  de l'élément  $\{v_{II}\}$  de  $\{v\}$ , dont le réflexe morphologique, que je ne tenterai pas d'analyser, est la gémation de  $C_2$ ; ce seul fait interdit la formation d'un PI, qui exigerait l'inclusion d'un deuxième élément,  $\{n_{pl}\}$ . De même des noms de métiers tel *xabbaazun* / *xabbaazuuna* 'boulangier(s)', si l'on y suppose l'inclusion d'un élément sémantiquement analogue au suffixe *-er* du français (ou de l'anglais *baker*) et distinct de  $\{n\}$ , peut-être additionné à lui, mais pas inclus (on limitera les inclusions à  $\{n\}$  aux éléments  $\{f\}$  et  $\{pl\}$ , plus l'ensemble vide).

Prenons en revanche le cas de *miftaaHun* / *mafaatiiHun* 'clef(s)'. À un certain stade de la langue, sa formation a dû impliquer la combinaison avec la racine  $R_{fH}$  'ouvrir' d'un élément ayant pour sens "Instrument" et dont le réflexe morphologique est le préfixe /mi-/. S'il en est ainsi, nous sommes dans une configuration d'addition, qui n'interdit pas la formation d'un PI aux termes du PIU. Mais le sens de l'élément "Instrument" suppose, pour que soit obtenue la signification finale de «clef», que la racine ait subi l'inclusion d'un élément de  $\{v\}$ . C'est cela qui rendrait impossible la formation d'un PI, n'était le processus de relexicalisation qui a consisté précisément à détacher *miftaaHun* de  $\{R_{fH} \{v\}\} = fataHa$  'il a ouvert', c'est-à-dire à annuler  $\{v\}$  dans l'ensemble total. Un PI peut dès lors être formé sans enfreindre le PIU<sup>69</sup>.

On doit maintenant poser la question : pourquoi une seule inclusion par racine ? La réponse est implicite dans la théorie des PI défendue ici : l'inclusion de  $\{n_{pl}\}$  active le site post- $C_2$ ; il n'y a qu'un seul site post- $C_2$  dans une racine, et un seul élément phonologique peut y être inséré. La conclusion s'impose<sup>70</sup>. Le PIU nous permet ainsi de rendre compte de la répartition des PI et des PE sans du tout faire appel à la distinction entre flexion et dérivation, et donc sans avoir à considérer les PE comme plus flexionnels que les PI.

## 6.2. L'épel morphologique

Les hypothèses précédentes sur la nature des PI et sur la constitution des ensembles lexicaux ou lexèmes me conduisent à développer une théorie *fonctionnelle* de l'épel. Pour les théories que Stump (2001) qualifie de *lexical-realizational*, telle la DM «classique», le vocabulaire est un répertoire d'objets phonologiques porteurs de traits, les morphèmes réalisés ou exposants (*exponents*), et l'épel consiste à les associer aux traits équivalents portés par les têtes syntaxiques fonctionnelles ou radicales. Les items du vocabulaire ou vocables sont donc ou bien des formes libres (par ex., turc *ev* 'maison'), ou bien des formes liées (par ex., turc /IVr/ 'pluriel'). Ces deux sortes de vocables ne sont pas différentes du point de vue de l'épel morphologique, mais seulement du point

de vue phonologique : les premiers sont prononçables isolément, les seconds, non ; les premiers ont une forme stable, les seconds peuvent varier dans les limites de processus phonologiques définis (cf. *evler* ‘maisons’ vs *başlar* ‘têtes’).

Les PI contredisent cette conception. Comme je crois l’avoir montré, la pluralité interne n’est pas (synchroniquement) associable à un morphème-objet ; la racine amplifiée s’oppose globalement à la racine simple ; les vocables qui réalisent l’une et l’autre sont des formes libres que l’on peut analyser, mais non pas en les segmentant en morphèmes-objets concaténés. Si l’on généralise cette observation, on en vient à la proposition suivante : le vocabulaire ne contient que des formes libres ; /IVr/ n’est pas un vocable, seuls *evler*, *başlar*, etc., en sont.

On est donc amené à concevoir l’épel comme une fonction qui prend pour domaine un ensemble d’éléments et lui fait correspondre un *objet morpho-phonologique* (MP – cf. Raffelsiefen 1992). Conformément aux hypothèses de la section 4, un objet MP est une suite ordonnée de positions telle que les positions impaires sont interprétées comme des attaques syllabiques (C), et les positions paires comme des noyaux (V), et chaque type est réalisé au moins une fois (cf. Lowenstamm 1996). C et V sont elles-mêmes associées à des segments phonologiques. La fonction épel sera donc notée  $f_{CV}$ . Soit  $\{e\}$  un ensemble d’éléments :

$$(30) \quad \{e\}_{f_{CV}} \rightarrow CV +$$

L’avantage de la solution fonctionnelle est qu’elle est capable de produire de la concaténation, comme je vais le montrer, mais qu’elle n’est pas obligée de le faire.

Prenons cet exemple typique de morphologie concaténative qu’est la forme turque *evler* ‘maisons’. Le lexème qui lui correspond est  $\{\{R_{ev}\}\{n\{pl\}\}\}$ ,  $\{pl\}$  inclus dans  $\{n\}$ , et l’ensemble  $\{n\{pl\}\}$  ou  $\{n_{pl}\}$  additionné à la racine. Cet ensemble lexical est le domaine de  $f_{CV}$ . Supposons que la fonction s’applique autant de fois que l’ensemble comprend de sous-ensembles combinés par la fonction lexicale d’addition, en commençant par le domaine de cette fonction, c’est-à-dire la racine. On définit ainsi un isomorphisme entre la fonction lexicale et la fonction épel. Ce sont donc ici deux applications, qui donnent les résultats suivants<sup>71</sup> :

$$(31) \quad \{R_{ev}\}_{f_{CV}} \rightarrow CV_e C_V V = /ev/$$

$$(32) \quad \{n\{pl\}\}_{f_{CV}} \rightarrow C_l V \sqcup C_r V = /IVr/$$

Ces deux applications sont composées :

$$(33) \quad < \{R_{ev}\}_{f_{CV}} \rightarrow CV_e C_V V \times \{n\{pl\}\}_{f_{CV}} \rightarrow C_l V \sqcup C_r V >_{f_{CV}} \rightarrow CV_e C_V V C_l V \sqcup C_r V = /evler/$$

(après Harmonie vocalique)

La linéarisation de l’objet MP résulte ainsi de la composition de la fonction épel réappliquée aux sous-ensembles délimités par la fonction addition.

Qu’en est-il quand la morphologie est « non concaténative » ? L’analyse des PI nous l’a montré : la non-concaténativité est l’expression morphologique de l’application de la fonction lexicale inclusion. Par exemple, l’ensemble



d'éléments lexicaux correspondant à *kataba* 'il a écrit' est quelque chose comme  $\{R_{ktb} \{v_{Tperf}\}\}$ <sup>72</sup>? Or, la fonction épel ne s'applique qu'aux sous-ensembles additionnés, elle ne « voit » pas les inclusions, si bien qu'elle devrait s'appliquer *in toto* à  $\{R_{ktb} \{v_{Tperf}\}\}$ , ce qui n'est pas le cas pour des raisons évidentes<sup>73</sup>. Je fais alors l'hypothèse suivante : la fonction épel  $f_{CV}$  peut se dédoubler en  $f_C$  et  $f_V$ .  $f_C$  a le même domaine que  $f_{CV}$ , à savoir le domaine de la fonction lexicale active, donc la racine ; elle lui fait correspondre une suite ordonnée de positions C (C +).  $f_V$  a pour domaine l'ensemble inclus, ici  $\{v_{Tperf}\}$  ; elle lui fait correspondre une suite ordonnée de positions V (V +). Ni C + ni V + ne sont des objets MP légitimes. Il s'ensuit que la composition de  $f_C$  et  $f_V$  a obligatoirement pour produit l'intercalation des V données par la seconde et des C données par la première<sup>74</sup> :

$$(34) \quad < \{R_{ktb} \{ \} \} \}_{f_C} \rightarrow C_k C_l C_b \times \{ \{v_{Tperf}\} \}_{f_V} \rightarrow V_a V_a >_{f_{CV}} \rightarrow C_k V_a C_l V_a C_b V_{\square} = /katab (a)/$$

Il n'y a plus dès lors d'isomorphie entre la linéarisation des segments de l'objet MP et la réitération de l'application fonctionnelle. Bref, on ne sait plus délimiter les morphèmes. Mais cela n'est pas un problème pour la théorie exposée ici. Comme je l'ai dit plus haut, elle peut produire de la concaténation, mais elle n'y est pas forcée, parce que le morphème comme objet n'appartient pas nécessairement à ses rouages<sup>75</sup>.

Considérons à présent la paire *kalbun* / *kilaabun* 'chien(s)', et d'abord l'épel du singulier. L'ensemble d'éléments que *kalbun* exprime, abstraction faite du Cas et du tanwîn, est  $\{R_{klb} \{n\}\}$ . La fonction épel composite  $f_C \times f_V$  s'y applique :

$$(35) \quad < \{R_{klb} \{ \} \} \}_{f_C} \rightarrow C_k C_l C_b \times \{ \{n\} \}_{f_V} \rightarrow V_a >_{f_{CV}} \rightarrow C_k V_a C_l V \{ \} C_b V = /kalb/$$

Cette représentation appelle quelques remarques. Premièrement, on voit que le site post- $C_2$  est suscité dans la suite CV terminale, qu'il soit ou non activé. La présence et la localisation du site sont en effet une conséquence nécessaire de la combinaison de la racine avec  $\{n\}$  (ou  $\{v\}$ , auquel cas, comme l'on sait, le site se fixe après  $C_1$ ).

D'autre part, (35) montre que la voyelle signature /a/ peut être considérée comme l'épel de  $\{n\}$  sans que cela nous contraigne à y voir, contre toute évidence, un morphème de nominalisation. Ce résultat est particulièrement intéressant pour l'analyse des noms verbaux tels que *qatlun* 'fait de tuer, meurtre' (cf. *qatala* 'tuer'). En effet, si l'ensemble lexical correspondant à *qatlun* est  $\{R_{qtl} \{v \{n\}\}\}$ ,  $\{n\}$  inclus dans  $\{v\}$ , et  $\{v \{n\}\}$  inclus dans l'ensemble radical, on voit que /a/ constitue en effet l'épel de  $\{v \{n\}\}$ , sans être pour autant un morphème nominalisateur analogue au suffixe – *at* d'*assassinat*, ce qu'il ne peut être. Peut-être doit-on même aller plus loin : dans la mesure où cette voyelle n'a pas de fonctionnalité, qu'elle n'entre dans aucune opposition, contrairement à la « mélodie »  $V_a V_a$  de (34) qui s'oppose à  $VV_u$  de *yaktubu* 'il écrit' et à  $V_u V_i$  de *kutiba* 'il a été écrit', il serait sans doute plus exact de supposer que  $f_V$  produit une voyelle non spécifiée  $V_{\square}$  au niveau MP où nous sommes, dont la valeur n'est déterminée qu'en phonologie. On expliquerait ainsi l'imprévisibilité de la voyelle



signature dans les noms verbaux ayant la forme d'un nom primitif (cf. *'ilmun* 'connaissance' associé à *'alima* 'savoir, connaître')<sup>76</sup>.

Enfin, l'épel de *kalbun* me justifie de ne pas avoir introduit la propagation autosegmentale dans celui de *kataba*. Car il faudrait alors mettre en place un dispositif spécial pour empêcher ici cette propagation aux positions V suivant C<sub>2</sub> et C<sub>3</sub>, et cela même dans l'hypothèse où la voyelle épelée par  $f_V$  est non spécifiée, car la suite CV terminale serait alors  $C_k V_{[]} C_1 V_{[]} \{ \} C_b V_{[]}$ , d'où *\*/kalaba/*<sup>77</sup>. Il est bien plus simple de considérer que  $f_V$  produit une seule position V, non spécifiée, qui s'intercale naturellement après la première position C. Les deux V suivantes restent non valuées. Je réécris donc (35) :

$$(36) \quad < \{R_{kib} \{ \} \} \}_{fc} \rightarrow C_k C_1 C_b \times \{ \{n\} \}_{f_V} \rightarrow V_{[]} \xrightarrow{f_{cv}} C_k V_{[]} C_1 V \{ \} C_b V = /kalb/$$

Il convient donc d'être plus précis quant à l'action des fonctions d'épel. En fait, les positions C et V sont données par le format CV, non par les fonctions. Ces dernières attribuent des valeurs, éventuellement non spécifiées, à ces positions, une valeur étant un ensemble ou une matrice d'éléments phonologiques que, pour simplifier, je représente par un phonème ou, pour les valeurs non spécifiées, par un vide entre crochets carrés. Une ou plusieurs positions peuvent être ignorées par les fonctions, auquel cas elles restent sans valeur, nues, phonologiquement invisibles. Une autre conséquence, sans doute de plus grande ampleur, est que  $f_V$  ne produit des valeurs spécifiées que pour les positions V dont l'épel a une fonctionnalité grammaticale. Toutes les autres V sont ou bien ignorées, ou bien pourvues de valeurs non spécifiées arbitrairement déterminées par la phonologie<sup>78</sup>. C'est en ce sens que la simple vocalisation d'une racine ne compte pas pour une opération morphologique.

Passons maintenant au PI *kilaabun*, épel de  $\{R_{kib} \{n_{pl}\}\}$ . La somme des hypothèses précédentes ne laisse qu'une solution :  $f_V$  s'applique à  $\{n_{pl}\}$  et lui fait correspondre un objet que je note V/C pour faire apparaître son ambivalence syllabique, c'est-à-dire un glide. Cette opération, la fonction  $f_V$ , et elle seule, peut l'accomplir, puisque, comme nous l'avons vu plus haut, la classe des glides est coextensive à la classe des voyelles, mais disjointe de la classe des consonnes.

L'objet V/C, n'étant ni C ni V, n'a pas sa place dans la suite CV +. Il doit donc s'insérer dans le site post-C<sub>2</sub>, dont la structure syllabique interne est indéterminée, comme on l'a vu. La fonction composite qui épelle *kilaabun* est ainsi la suivante<sup>79</sup> :

$$(37) \quad < \{R_{kib} \{ \} \} \}_{fc} \rightarrow C_k C_1 C_b \times \{ \{n_{pl}\} \}_{f_V} \rightarrow V/C_A \xrightarrow{f_{cv}} C_k V C_1 V \{C_A V\} C_b V = /kilaab/$$

Cette séquence d'opérations a une conséquence intéressante et, pour tout dire, heureuse : les voyelles ne sont pas produites par la fonction épel  $f_V$ , puisque celle-ci ne donne que V/C. Elles doivent donc l'être par des opérations phonologiques ultérieures, en accord avec nos conclusions précédentes : la voyelle longue /aa/ provient de la propagation de /A/ aux positions V adjacentes, accompagnée de l'effacement du glide ; /i/ est épenthétique.

La fonction épel s'applique de même quant aux autres types de PI formés sur des racines 3C (*najmun* / *nujuumun*, *madiinatun* / *mudunun*, *ḡurfatun* / *ḡurafun*, etc.) ou 4C (*jundubun* / *janaadibun*). 'ville(s)'. Il convient toutefois d'examiner son application aux racines 2C (*damun* / *dimaa'un*) et aux cas plus complexes (*diibaajun* / *dabaabiiju* ~ *dayaabiiju*).

Le singulier *damun* ne semble pas poser de problèmes particuliers. En première analyse, il paraît suffisant de supposer que  $f_C$  produit deux consonnes au lieu de trois, d'où résulte la suite CV terminale  $C_d V_a C_m V \{ \}$ . Lorsque  $\{n_{pl}\}$  est dans l'ensemble lexical, en revanche, non seulement le site post- $C_2$  est activé, mais la condition du triconsonantisme des racines morphologiquement traités prend effet. En fait, ces deux phénomènes sont très probablement connexes. La proposition suivante suffit à le montrer : le site activé doit être *interne*, c'est-à-dire non seulement post- $C_2$ , mais aussi pré- $C_3$ . Le dégagement d'une troisième consonne est une conséquence immédiate de l'activation du site. Il ne serait toutefois guère cohérent de supposer que  $f_C$  produit  $C_3$  lorsque  $\{n_{pl}\}$  est inclus dans l'ensemble lexical, pour la bonne raison que  $\{n_{pl}\}$  n'est pas le domaine de cette fonction ; et aussi parce que, comme nous venons de le voir, les fonctions d'épel ne produisent pas, mais donnent des valeurs aux positions, celles-ci étant indépendamment fournies par le format CV. On considérera donc plutôt que  $f_C$ , en tant que fonction morphologique, value *toujours* au moins trois C, mais la troisième peut ne pas être interprétée et rester latente si le site est inactif (voir plus loin). L'épel de *damun* se fait donc comme suit :

$$(38) < \{R_{dm} \{ \} \} f_C \rightarrow C_d C_m C_{\square} \times \{ \{n\} \} f_V \rightarrow V_{\square} >_{f_{cv}} \rightarrow C_d V_{\square} C_m V \{ \} C_{\square} V = /dam/$$

Lorsque le site est activé par  $\{n_{pl}\}$ , en revanche, la consonne latente reçoit la valeur par défaut occlusive glottale (OG) dans la suite CV terminale, afin de satisfaire la propriété d'intériorité du site activé :

$$(39) < \{R_{dm} \{ \} \} f_C \rightarrow C_d C_m C_{\square} \times \{ \{n_{pl}\} \} f_V \rightarrow V / C_A >_{f_{cv}} \rightarrow C_d V C_m V \{ C_A V \} C_{OG} V = /dimaa' /$$

La suite CV terminale passe ensuite par les mêmes processus phonologiques que *kilaabun*. Pour le dire autrement, et dans le souci d'éviter tout malentendu, la racine de *damun* est réellement, lexicalement, 2C ; la troisième consonne est le produit de l'application de la fonction morphologique  $f_C$ , et elle n'est réalisée que si l'activation du site post- $C_2$  exige qu'elle le soit.

Voyons maintenant *diibaajun* / *dibaabiiju* ~ *dayaabiiju* 'brocart(s)'. Nous avons vu que la racine peut être considérée comme effectivement 5C, deux des radicales étant des glides :  $R_{dlbAj}$ . La vocalisation du singulier est donc totalement dérivable de l'hypothèse CV, puisque les deux voyelles longues résultent de la propagation des glides /I/ et /A/ en C aux V adjacentes, si bien que la fonction  $f_V$  s'applique à vide, ce que je note par  $f_V \rightarrow V$  :

$$(40) < \{R_{dlbAj} \{ \} \} f_C \rightarrow C_d C_i C_b C_A C_j \times \{ \{n\} \} f_V \rightarrow V >_{f_{cv}} \rightarrow C_d V C_i V \{ \} C_b V C_A V C_j V = /diibaaj/$$

Le nombre des positions C valuées par  $f_C$  n'est pas limité vers le haut quand le site est inactif. Quand le site est actif, en revanche, sa borne supérieure est 4<sup>80</sup>. Une des C radicales de (40) doit donc recevoir de  $f_{CV}$  une valeur non spécifiée (d'où sa notation par C  $\square$  dans la suite CV terminale ci-dessous). Une C non spécifiée, plus généralement, une position non spécifiée doit être identifiée par une autre position qui la gouverne, comme on l'a vu, à moins qu'elle ne soit en situation d'être réalisée par défaut (voir ci-dessous). Si  $C_2 = /I/$  est affectée, le PI est *dabaabijju*, et son épel procède ainsi :

$$(41) \quad < \{R_{dlbA_j} \{ \} \} \}_{f_C} \rightarrow C_d C_1 C_b C_A C_j \times \{ \{n\} \}_{f_V} \rightarrow V/C_A >_{f_{CV}} \rightarrow C_d C_{\square} V \{C_A V\} C_b V C_A V C_j V = /daC_{\square} aabVAVj/$$

Ce résultat est phonologiquement mal formé : la position non spécifiée doit être interprétée, car, si elle ne l'était pas, il s'ensuivrait un PI \*/daabijj/ qui serait quantitativement égal au singulier ; elle l'est pas copie de  $C_3$ , c'est-à-dire de la position gouvernante, d'où /dabaabVAVj/ = /dabaabijj/ après ancrage de /I/ épenthétique à la position occupée par /A/ radical au singulier<sup>81</sup>. Au PI alternatif *dayaabijju*, c'est  $C_4 = /A/$  qui est laissée non spécifiée par  $f_{CV}$ .

On aura sans doute remarqué que la suite CV terminale de (41) paraît enfreindre la condition du maximum de quatre consonnes (condition 4C), puisqu'elle en contient cinq :  $C_d$ , C  $\square$ ,  $C_b$ ,  $C_A$  et  $C_j$ , et de même  $C_d$ ,  $C_1$ ,  $C_b$ , C  $\square$  et  $C_j$  dans le PI alternatif *dayaabijju*. La forme de surface est pourtant licite, du fait que, dans l'une et l'autre forme,  $C_4$  a donné naissance à une voyelle longue, si bien qu'il n'apparaît en effet que quatre consonnes.

Cette observation nous conduit à une conclusion intéressante, à savoir que l'élagage total d'une consonne afin de satisfaire les conditions de taille de la racine constitue une solution de dernier recours. C'est la seule solution avec les noms de racine 5C tels que *ankabuuTun*, car si l'une des consonnes radicales, disons /T/, était seulement non spécifiée dans l'épel du PI, elle devrait être identifiée par une autre radicale, en sorte que le nombre final de consonnes ne changerait pas. En outre, si l'identification se fait depuis une position gouvernante, une radicale finale non spécifiée ne peut être identifiée puisqu'elle n'a aucune radicale à sa droite. C'est précisément ce qui se passe au singulier *damun* :  $C_3$  non spécifiée n'est pas identifiée, ce qui est sans conséquence puisque aucune condition de taille ne s'impose au singulier morphologiquement non traité. Au pluriel (*dimaa'un*),  $C_3$  n'est toujours pas identifiable par gouvernement, mais elle doit recevoir une valeur spécifiée, qui lui vient par défaut. Cette solution est en principe applicable à la consonne finale de *ankabuuTun*, mais pas en pratique car la condition 4C doit être satisfaite. Il ne reste donc qu'à ne pas réaliser du tout cette consonne, étant donné que *ankabuuTun* ne contient par ailleurs aucun segment faible, glide ou sonante, qui pourrait être affecté de préférence à  $C_5$ <sup>82</sup>.

Considérons d'autre part ce qui se passerait si le /A/ radical de *diibaajun* était réalisé par défaut ou identifié par gouvernement au PI. Selon que /I/ radical est lui-même valué ou identifié par gouvernement, cela nous donnerait les formes suivantes : \*/dayabaa'ij/ ou \*/dabaaba'ij/ (cf. *šajaa'iru* 'cigarettes'), \*/dayaabajj/

ou \*/dabaabajj/. Toutes enfreignent la condition 4C. L'échappatoire consiste donc, on l'a vu, à ancrer le /l/ épenthétique à la position occupée par /A/, produisant ainsi des formes *dabaabiiju* et *dayaabiiju* qui ne contreviennent pas apparemment à la condition 4C, ce qui suffit à les rendre légitimes. Et, puisque l'échappatoire existe, elle doit être adoptée vu le caractère de dernier recours de l'élagage total. On comprend ainsi que ne soient pas produites des formes théoriquement possibles telles que \*/dabaabij/ ou \*/dayaabij/, avec un voyelle épenthétique brève due à la suppression complète de /A/.

Nous verrons bientôt un autre exemple des conséquences de ce caractère de dernier recours de l'élagage. Il me paraît important de souligner ici que le raisonnement ci-dessus, qui nous a permis de rendre compte tout à fait de ces formations complexes, est rendu possible par l'hypothèse que les voyelles longues trahissent toujours la présence de glides sous-jacents, hypothèse qui constitue l'autre pilier de la présente théorie à côté de et en liaison avec l'existence du site post-C<sub>2</sub><sup>83</sup>.

## 7. Conclusion

Les PI illustrent la thèse saussurienne de la différence comme principe organisateur du langage : le contraste entre la racine simple interprétée au singulier et la racine amplifiée interprétée au pluriel est une « pure » différence, en cela qu'elle ne dépend pas de la présence absence d'un morphème-objet auquel la signification « pluriel » serait spécifiquement attachée. Par conséquent, les PI ne sont pas formés à partir des singuliers correspondants ; PI et singuliers constituent deux épels parallèles depuis deux ensembles d'éléments partageant une même racine. Là est l'idée première de cet article.

Une objection qu'on peut lui faire est que, parfois, la forme du PI semble dépendre de celle du singulier. Le cas le plus net est celui des noms à racine 4C + : si le singulier comporte une voyelle longue après C<sub>3</sub>, le PI en comporte une dans la même position (cf. *miftaaHun* / *mafaatiiHu* 'clef(s)', *finjaanun* / *fanaajiiinu* 'tasse(s)', etc.) ; sinon, le PI présente une voyelle brève (cf. *manzilun* / *manaazilu* 'habitation(s)', *jawharun* / *jawaahiru* 'bijou(x)', etc.). La durée de la voyelle du singulier paraît donc avoir été transférée à la voyelle épenthétique correspondante du PI, ce qui laisse penser que celui-ci est bien formé à partir de celui-là. Une autre analyse est néanmoins possible.

À l'origine, on le sait, *miftaaHun* est un nom d'instrument formé sur la racine *ftH*. A ce stade, la suite CV terminale est :

$$(42) \quad mi-C_fVC_fV\{C_AV\}C_HV = /mi-ftaaH/$$

où /mi/ est un préfixe produit par l'application de la fonction  $f_{CV}$  non composite également disponible en arabe (voir plus loin) à un élément lexical voulant dire « instrument ». L'important est que la dérivation instrumentale met en jeu l'activation du site post-C<sub>2</sub> /t/. Après la relexicalisation et la réanalyse du /m/ préfixal comme C<sub>1</sub>, la suite CV devient :

$$(43) \quad C_mVC_fV\{C_tVC_AVC_HV = /miftaaH/$$

Le site se repositionne après la nouvelle  $C_2$  /f/, et le glide /A/ source de la voyelle longue est lui aussi réanalysé comme une radicale,  $C_4$ . La racine devient ainsi une racine 5C. Notez que la voyelle suivant  $C_1$  /m/ ne reçoit pas de valeur de  $f_V$ , car elle est devenue une voyelle épenthétique destinée à briser le groupe consonantique initial \*/mf/. La forme *mafaatiiHu* du PI est dès lors prévisible:  $C_4$  /A/ est dévaluée en vertu de la loi régulant le format maximal de la racine utile, tandis que le site est activé par insertion de /A/ en C. D'où la suite CV terminale:

$$(44) \quad C_mVC_fV\{C_AV\}C_tVC_{\square}VC_HV = /mafaatVC_{\square}VH/$$

En s'attachant à la position C non valuée, /I/ produit automatiquement une voyelle longue /i:/ (cf. ci-dessus l'épel de *dayaabiiju*), si bien que la forme apparente n'enfreint pas la condition 4C. Là encore, l'existence de cette solution écarte d'emblée la solution de dernier recours qui consisterait à élaguer  $C_5$ , ce qui obligerait à réaliser /A/ par défaut, d'où \*/mafaati' /.

Soit maintenant *jawharun* / *jawaahiru* 'bijou(x)'. Il s'agit là d'une racine 4C originelle, avec au singulier la suite CV suivante:

$$(45) \quad C_jV_{\square}C_UV\{C_hV_{\square}C_rV = /jawhar/$$

Au PI, le segment épenthétique se lie tout naturellement à la voyelle non spécifiée entre  $C_3$  et  $C_4$ , d'où pour *jawaahiru* la suite CV terminale suivante où  $V_{[i]} = V_{\square}$  phonologiquement interprétée comme /i/:

$$(46) \quad C_jV_{\square}C_UV\{C_AV\}C_hV_{[i]}C_rV = /jawaahir/$$

Il n'est donc pas nécessaire de supposer l'existence d'une connexion transversale entre les épels parallèles du singulier et du PI. Les principes d'épel eux-mêmes et la condition 4C suffisent à produire la correspondance entre les singuliers CCCvVC et les PI CaCaaCiiC. Quant à cette condition, il importe de noter qu'elle n'institue pas une correspondance entre singuliers et pluriels particuliers, mais entre racines simples et racines utiles: si une racine simple a cinq consonnes ou plus, alors la racine utile doit en avoir quatre, c'est-à-dire le nombre autorisé le plus proche, faisant que la racine utile diffère au minimum de la racine simple (on peut voir là une sorte de principe d'économie). On sait du reste que la condition s'applique aussi à la formation des diminutifs (cf. '*andaliibun* 'rossignol' / '*unaydilun* 'petit rossignol').

Il ne m'est pas possible d'examiner ici les autres cas pour lesquels une telle connexion est supposée indispensable<sup>84</sup>. Je les crois réductibles de la même manière, et j'espère pouvoir le montrer dans un travail plus développé.

À la fin de la section 3, j'ai noté deux faits d'observation en promettant de les interpréter plus tard. L'un est que les noms ayant un gabarit CVCC tels que *kalbun* ne peuvent être des PI. Il est désormais superflu de démontrer que cette propriété négative découle directement de la formation des PI telle que je l'analyse. En effet, puisque cette formation met essentiellement en jeu

l'occupation du site post- $C_2$  par un glide qui se manifeste sous la forme d'une voyelle longue ou brève, il n'y a de fait aucune possibilité pour que CVCC soit un gabarit de PI. C'est là quelque chose que le modèle de McCarthy & Prince ne prédit pas: forcés qu'ils sont d'accepter l'existence d'autres prosodies que l'iambe, ils ne peuvent exclure la possibilité théorique de former un PI en conservant la prosodie du singulier et en changeant la qualité de la voyelle selon un principe de polarité, si bien que \*/kilb/, par exemple, serait le PI de *kalbun*. La théorie ici défendue rend une telle formation intrinsèquement impossible, comme les faits l'exigent.

L'autre fait d'observation est que les PI sont toujours quantitativement différents des singuliers correspondants, c'est-à-dire qu'ils comptent plus ou moins de mores, mais jamais autant<sup>85</sup>. Quoique je n'ai pas ici la place de le démontrer complètement, c'est là encore une conséquence de l'occupation du site par un glide en position C ou V. Quelques exemples le montreront suffisamment: /ka<sup>1</sup>lV<sup>1</sup>{ }bV<sup>1</sup>/ = 3 mores vs. /ki<sup>1</sup>l{aa}<sup>2</sup>bV<sup>1</sup>/ = 4 mores; /ki<sup>1</sup>t{aa}<sup>2</sup>bV<sup>1</sup>/ = 4 mores vs /ku<sup>1</sup>t{u}<sup>1</sup>bV<sup>1</sup>/ = 3 mores; /ġu<sup>1</sup>rV<sup>1</sup>{ }fa<sup>1</sup>tV<sup>1</sup>/ = 4 mores vs /ġu<sup>1</sup>r{a}<sup>1</sup>fV<sup>1</sup>/ = 3 mores (les chiffres en exposant indiquent le nombre de mores en même temps que les frontières de syllabes dans l'hypothèse CV).

Une autre thèse défendue dans cet article est que la non-concaténativité de la morphologie de l'arabe, la morphologie non concaténative en général, peut être expliquée comme découlant du dédoublement de la fonction épel  $f_{CV}$  en une fonction  $f_C$  et une fonction  $f_V$  composées. Appelons ce mode de l'épel le mode composite. Les morphologies concaténatives, celle du turc par exemple, manifestent pour leur part le mode non composite ou direct de l'épel. On se tromperait, néanmoins, si l'on considérait ces deux modes comme deux caractéristiques typologiques associées à des langues particulières. Il s'agit plutôt des deux manières possibles, les deux seules, me semble-t-il, qu'a la fonction épel de s'appliquer. L'anglais, par exemple, où le mode direct est largement prédominant, n'ignore pas le mode composite, comme le montrent les pluriels et les prétérits apophoniques. En arabe, il paraît clair que tous les mots que la tradition grammaticale range sous l'appellation de «particules» participent de l'épel direct (cf. *'ilaa* 'vers', *fii* 'dans', *wa* 'et', etc.). On constate également que les deux modes d'épels peuvent se combiner dans les limites d'un même mot: ainsi des noms préfixés où le préfixe n'est pas intégré à la racine et qui, de manière symptomatique, se refusent à la pluralisation interne (par exemple, *mu<sup>c</sup>allimun* / *mu<sup>c</sup>allimuuna*); ainsi encore des préfixes et suffixes qui épellent les traits de personne, de genre et de nombre des verbes conjugués (cf. *katabna* 'nous avons écrit', *yakubu* 'il écrit', etc.). Bref, la (non-)concaténativité, loin d'être un primitif, ne serait que la manifestation superficielle des deux modes d'épel possibles, auxquels toute langue peut avoir recours.

## NOTES

1. Par « arabe », on comprendra l'arabe littéral moderne (*Modern Standard Arabic*). Je ne dirai rien des dialectes, qui, du reste, ont conservé le système classique à peu près intact. (Cela est vrai même d'un dialecte aussi divergent que le maltais.) Afin de ne pas allonger démesurément cet article, je ne traiterai que des noms.
2. Dans les langues sans gabarits déterminés, les seules positions toujours repérables sont le début et la fin du mot, ce pour quoi sans doute ces langues font peu d'usage morphologique de l'intérieur du mot. Même alors, pourtant, la position n'est jamais un critère suffisant : par exemple, le /s/ final du portugais *lápís* 'crayon' n'est pas un morphème malgré sa forme et sa position. Dans une langue gabaritique, en revanche, l'occurrence de matériel phonologique dans une position de site en signale sans équivoque la nature morphologique.
3. Au nominatif ; à l'oblique (accusatif et génitif) les suffixes sont /-iina/ et /-aatin/. Par « base », j'entends la forme produite par l'application d'un gabarit à la racine consonantique, incluant éventuellement un préfixe, mais à l'exclusion de tout suffixe, par exemple, /muʔallim/. Les exemples seront toujours cités au nominatif et avec le /n/ final nommé *tanwīn* (la forme « indéfinie »). On sait que les voyelles casuelles et le *tanwīn*, ainsi que le /t/ de féminin (*taʔmarbuuTa*) ne se prononcent pas à la pause, voire pas du tout. Je ne l'indique pas dans la translittération. Cette présentation de /-uuna/ et de /-aatun/ comme des suffixes inanalysables paraît justifiée d'un point de vue synchronique. La réalité historique est peut-être différente (cf. Ratcliffe 1998).
4. Le PE masculin s'applique en outre aux noms propres masculins, tandis que le PE féminin s'applique aux noms propres féminins, aux diminutifs, aux noms des lettres de l'alphabet (par exemple, *Taa'un* / *Taa'aatun* 'Ta(s)'), aux noms des mois et aux emprunts non intégrés (par exemple, *baaraašuuTu* / *baaraašuuTaatun* 'parachute(s)').
5. Cf. Ratcliffe (1998: 224-229) quant aux origines possibles de cet élément qui ne peut pas être analysé comme un morphème.
6. Pour mieux faire ressortir le procédé, les consonnes radicales sont notées en majuscules. (1) n'est qu'un échantillon, qui ignore en particulier les formations mettant en jeu une modification à droite de la racine (par exemple, *waziirun* / *wuzaraa'u* 'ministre(s)', *baladun* / *buldaanun* 'pays', *baabun* / *bibaanun* 'porte(s)', *gazaalun* / *gizlaanun* 'gazelle(s)', *diikun* / *diyakatun* 'coq(s)', etc.). Faute de place, celles-ci ne seront pas traitées dans cet article, quoiqu'elles soient tout à fait compatibles avec la théorie qui y sera développée.
7. Les PI des 4C + se signalent en outre par le fait d'être des diptotes, c'est-à-dire de ne distinguer que deux cas, nominatif et oblique, et de ne pas comporter le *tanwīn* à la forme indéfinie. Je n'ai rien à dire de cette propriété.
8. On sait que, pour l'accord, les PI sont féminins singuliers, sauf s'ils dénotent des humains mâles, auquel cas l'accord se fait au masculin pluriel. (Mais l'accord au féminin pluriel est admis et tend à devenir la norme quand le PI dénote des femmes.) Je reviendrai sur ce point.
9. Il est possible que la pluralisation soit, d'une façon générale, un phénomène dérivationnel (cf. Beard 1995). Je ne m'engage pas sur ce point, tout comme je laisse de côté les difficultés que soulève, on le sait, toute tentative pour séparer à coup sûr la dérivation de la flexion (cf. Scalise 1984 ; Dressler 1989). Rien de tout cela n'altère le fait que les PI arabes se situent très clairement dans le camp de la dérivation.
10. La conjonction des deux propriétés est décisive : les formations flexionnelles peuvent être multiples, elles ne sont qu'exceptionnellement imprévisibles, dès lors que les caractères de la base (classe morphologique, genre, etc.) sont connus.

11. Souvent, les variantes, quand elles sont également acceptables, tendent à se spécialiser ou bien le sont d'emblée : cf. anglais *mice* 'des souris' vs *mouses* 'des souris d'ordinateur'. Leur coexistence est généralement le résultat d'un mélange de dialectes, et tel est sans doute le cas en arabe. La particularité (relative) de celui-ci est de les avoir toutes conservées, sans doute dans le but délibéré d'accroître les ressources littéraires et poétiques de la langue. Il n'y a rien là d'extraordinaire en soi, mais cela confirme, et c'est ce qui nous importe, la nature dérivationnelle des PI, car la flexion, temporelle ou casuelle, par exemple, n'offre pas ces ressources.

12. Ce caractère dérivationnel est en fait reconnu, au moins implicitement, depuis longtemps, et c'est lui qui a motivé les tentatives récurrentes pour démontrer que les PI ne sont pas vraiment des pluriels, qu'ils sont « moins pluriels » que les PE, et pour les rapprocher de formations comme les collectifs ou les noms verbaux ou *maṣdar* (cf., par exemple, Barth 1904, Petráček 1960, Corriente 1971). Je n'entre pas dans cette discussion historique, et très spéculative, considérant que, du point de vue synchronique, les PI présentent tous les emplois de vrais pluriels. Autrement dit, *mulūkun*, PI de *malikun*, ne dénote pas moins une pluralité dénombrable de rois que *mu'allimūna*, PE de *mu'allimun*, ne dénote une pluralité dénombrable de professeurs. Pour la même raison, je ne dirai rien des « pluriels de paucité » (par exemple, *farxun* / 'afraaxun 'oisillon(s)', *ṣayxun* / ṣiixatun 'vieillard(s)') dont il n'est plus vrai, si ce le fut jamais, qu'ils ne valent que de quantités inférieures à dix (cf. Blachère & Gaudetroy-Demombynes 1975 : 178-182).

13. Il convient de distinguer la signification (*Bedeutung* dans la tradition frégéenne), qui implique une dénotation, du sens (*Sinn*) qui ne l'implique pas. En toute rigueur, une racine n'a pas de signification, elle n'a qu'un sens générique. En pratique, et en arabe, on peut évaluer le sens d'une racine à la signification du verbe de mesure I ou du nom primitif qui la réalise.

14. Autrement dit, la racine de *miftaahun* est *mftH* et non *ftH*. On citera encore le cas des noms verbaux comme *tajribatun* (cf. *jarraba* 'essayer', mesure II) qui forment un PI *tajaaribu* au sens d'expériences scientifiques, et un PE *tajribaatun* au sens d'essais, faits d'essayer. Signalons toutefois que les noms verbaux des verbes de mesure I ne forment que des PI, à l'instar des noms primitifs dont ils ont l'aspect (cf. *'ilmun* / *'ulumun* 'savoir(s)', science(s)', nom verbal de *'alima* 'savoir').

15. En fait, Ratcliffe considère que la morphologie dérivationnelle opère sur des lexèmes au sens d'Aronoff (1994), et la racine est pour lui une abstraction. D'où la nécessité de séparer en deux niveaux les lexèmes « terminés », qui ne peuvent qu'être prolongés par des affixes, et les lexèmes modifiables de l'intérieur.

16. Cela en supposant, comme le font les grammairiens, que *'ayaadin* est formé sur la pseudo racine *'ydn* selon le gabarit CaCaCiC des 4C +, plutôt que sur la racine *yd* parallèlement à *'aydin*. Le contraste avec *Huruubaatun* subsiste même dans cette dernière hypothèse, mais ses conséquences sont moins intéressantes. Arad (ce volume) étudie un phénomène du même ordre en hébreu moderne.

17. En réalité, on verra plus loin qu'il s'agit plus probablement d'une racine 3C dont la deuxième et la troisième radicale sont des glides sous spécifiés.

18. De même pour le diminutif : *'unaykibun* 'petite araignée'. Les racines supérieures à 5C sont rares. Voici un exemple de nom à racine 6C, qui doit donc perdre deux consonnes pour former son PI : *'imbaraaTiurun* / *'abaaTiratun* 'empereur(s)'. Les consonnes qui tombent sont ici C<sub>2</sub> /m/ et C<sub>4</sub> /r/, soit deux sonantes, plutôt que C<sub>5</sub> /T/ et C<sub>6</sub> /r/. Cela est significatif, et j'y reviendrai.



19. Il y a chez Derenbourg bien des aperçus fascinants, qui ont contribué à inspirer les analyses qui seront développées plus avant.
20. La thèse de Levy (1971) est une tentative, dont l'auteur reconnaît elle-même qu'elle est peu satisfaisante, pour rendre compte des PI dans le cadre linéaire de Chomsky & Halle (1968). La valeur du travail réside principalement dans l'étude très détaillée qu'elle contient des diverses formes de PI et leur corrélation statistique avec les singuliers correspondants. L'article de McCarthy & Prince est à la fois un prolongement et une réponse à un article antérieur de Hammond (1988). Ils n'ont pu manquer d'être inspirés par le travail de Greenberg (1955), qui met en valeur la fréquence d'un pluriel interne en /a/ dans les langues afroasiatiques. Tous ces travaux, sauf Asfour (2001), ont pour point commun de présupposer que les PI sont dérivés des singuliers, thèse avec laquelle j'exprimerai plus loin mon désaccord. Les thèses d'Asfour, dont j'ai déjà dit qu'elles sont proches des miennes, seront commentées à la section suivante. Parmi les études non génératives, il convient de citer encore la série d'articles de Petráček (1960) sur la flexion interne, ainsi que Murtonen (1964) et Corriente (1971). Pour les PI en dehors de l'arabe, on citera Angoujard & Denais (1989).
21. L'analyse défendue par McCarthy dans sa thèse de 1979, publiée en 1982, peut en effet être considérée comme rendue obsolète par celle de son article commun avec Prince.
22. L'exemple de *jabalun* montre qu'il ne s'agit pas de syllabes de surface, mais de la syllabe lourde que l'on forme en circonscrivant les positions CVC initiales du squelette.
23. La voyelle brève n'appartient pas à l'affixe et doit être considérée comme épenthétique (voir plus loin).
24. En fait, le vrai nom du pied brève – brève, pour autant que c'en est un, est «pyrrhique». Un trochée est le contraire d'un iambe, soit longue – brève, si bien que le terme ne s'applique vraiment qu'aux items comme *'anjumun* (syllabifié /'an-ju-mun/).
25. En revanche, et paradoxalement pour la thèse de McCarthy & Prince, la séquence de deux /u/ brefs de *kutubun* 'livres' paraît être exclusivement associée au pluriel pour les 3C. Seuls les 4C la présentent associée au singulier (cf. *funduqun* 'hôtel').
26. Cf. l'analyse de la reduplication de Marantz (1982), sur laquelle McCarthy & Prince s'appuient.
27. La notion même d'activité morphologique n'a pas de sens dans une théorie qui ne reconnaît que des morphèmes-objets et leur affixation.
28. Les 22 exceptions apparentes sont en fait des singuliers collectifs (cf. *rakbun* '(troupe de) cavaliers, cavalcade' vs *raakib<sup>um</sup>* / *rukkaab<sup>um</sup>* 'cavalier(s)').
29. On admettra que la consonne finale compte pour une more, étant suivie d'une position vocalique virtuelle.
30. Les étymons de Bohas sont presque tous 2C et leurs consonnes ne sont pas, ou pas complètement ordonnées, hypothèse qui n'a pas d'incidence sur mon propos. Cette question dite du *Biradikalismus* des racines sémitiques a fait l'objet d'un vaste débat au XIX<sup>e</sup> siècle. Bohas, qui a eu le courage de la reprendre, est l'un des premiers à avancer des arguments convaincants en sa faveur. S'il a raison, cela contribue à expliquer la position du site. S'il a tort, l'hypothèse n'en est pas affectée.
31. S'agissant des racines nominalisées (voir plus loin), la seule opération susceptible de les affecter est la pluralisation interne. La pluralisation externe (y compris la formation du duel) et la flexion casuelle sont des opérations extérieures à la racine, qui ne la touchent pas.
32. Le noyau peut être complexe, auquel cas on a une diphtongue, mais il n'existe pas de codas.

33. Inversement, du fait de l'asymétrie, ce qui se lie à V peut se lier à C. Voilà qui donne à penser qu'il doit être possible de simplifier encore l'analyse, comme me le suggère Jean Lowenstamm (c.p.). Ce sera pour un travail futur.

34. Quant à la qualité, on sait toutefois que la probabilité des trois éléments n'est pas égale : /A/ et /U/ sont les plus fréquents, avec peut-être une certaine prédominance de /A/, alors que /I/ est rare (cf. *kaliibun*, variante de *kilaabun* 'chiens'). L'explication est sans doute historique.

35. J'utilise le français comme métalangue. Un symbole indépendant de toute langue particulière serait préférable, mais peu commode. Par « ensemble d'éléments », j'entends à peu près la même chose que le *feature bundle* de la DM (voir section suivante).

36. C'est ici qu'il peut y avoir confusion entre la racine comme lexème abstrait et la racine comme suite de consonnes. On notera que cette ambiguïté n'est pas propre à notre objet, ainsi lorsqu'on dit que *confondre* et *confusion* ont la même racine. Il devrait suffire de savoir à quel niveau on raisonne pour l'éviter.

37. Une particularité de la phonologie de l'arabe est que le glide, n'a pas besoin d'être identique aux voyelles qui l'encadrent pour s'effacer : /aC<sub>1</sub>a/ donne également /a:/. Dans le cas présent, néanmoins, nous savons que le glide ne peut être que /A/, vu la valeur de la voyelle longue (/C<sub>U</sub>/ donnerait /u:/ et /C<sub>I</sub>/, /i:/).

38. Dans un système à trois voyelles comme celui de l'arabe littéral, /I/ est [+haut], /U/ est [+haut, +arrière], et /A/ est spécifié négativement pour tous les traits.

39. Il s'agit d'une représentation allégée. Cf. section suivante.

40. Le fait que *kaatibun* a deux PI selon sa signification démontre encore une fois que les PI ne sont pas dérivés des singuliers, mais constituent des formations parallèles.

41. Notez que CaCaCatun est aussi un gabarit de PI (cf. *Talabatun* 'étudiants'). Il s'agit donc peut-être d'une repluralisation.

42. On pourrait aussi considérer *ğurafun* comme un collectif dont *ğurfatun* serait le singulatif, sur le modèle de *šajarun* 'des arbres' vs *šajaratun* 'un arbre'. Il se peut que cette explication ait valeur historique, mais il n'en reste pas moins que *ğurafun* est sémantiquement un pluriel et non un collectif (cf. *xamsu ġurafin* 'cinq chambres').

43. Vu l'équivalence de l'occlusive glottale et de /A/, une représentation plus exacte est /d.m.{C<sub>A</sub>.}C<sub>A</sub>/. On notera que ces deux segments identiques successifs ne peuvent être confondus par le principe du contour obligatoire, car, au niveau phonologique où celui-ci s'applique, ils sont distincts.

44. Cf. la note précédente. Une autre analyse possible serait que la racine est 2C, et l'occlusive glottale un simple anti-hiatus phonétique. L'analyse proposée dans le texte me paraît supérieure, car elle prédit la forme du PI.

45. Lorsque C<sub>5</sub> est élagué, en revanche, la consonne disparaît sans laisser de trace (cf. *ʿankabuuTun* / *ʿanaakibu* 'araignée(s)').

46. C'est à ce niveau que les vocables acquièrent la totalité de leurs significations encyclopédiques, au-delà de la dénotation – par exemple, le fait que « incendie » se dit d'une maison qui brûle, pas d'un papier qui brûle.

47. Ce qui ne veut pas dire que cette projection n'existe pas, mais seulement que mon analyse de la formation des PI n'en a pas besoin.

48. Un autre raison pour considérer que les ensembles d'éléments doivent être présents comme tels dans la syntaxe est l'accord, car les éléments pertinents de la source doivent être accessibles à la cible, et cela même si l'accord lui-même ne met pas nécessairement en jeu un déplacement syntaxique, visible ou non, de la cible vers la source (cf. Chomsky 1999).

49. Notez que l'exigence de distinctivité n'implique pas que le morphème a nécessairement une forme phonologique (« zéro » peut être distinctif).
50. Cf. Kayne (1994) quant aux raisons pour lesquelles l'adjonction doit se faire à gauche.
51. Sur le fait que les PI arabes ne sont *pas* apophoniques, cf. Ségéral (1995).
52. On pourrait dire, sans trop pousser les choses, que  $R_{klb}$  est ce que nous avons à l'esprit quand nous avons le mot *chien* "sur le bout de la langue".
53. Quoique catégoriellement non spécifiées, il se pourrait que les racines soient *a priori* classées comme dénotant des événements ou des états (cf. Nash 2002).
54. Je répète que « destruc » est une pure étiquette identifiant une racine abstraite dont la réalité est sémantique. Il ne s'agit en aucune façon d'une *forme* sous-jacente à *détruire* et à *destruction*.
55. Je reste prudemment agnostique quant aux types d'êtres logico-sémantiques (cognitifs) dont ces éléments représentent la traduction linguistique. On sait que  $v$  (« petit  $v$  ») joue un rôle dans la théorie minimaliste en tant que catégorie fonctionnelle dominant  $V$  et lui attribuant la transitivité en même temps qu'elle légitime l'argument externe (cf. Kratzer 1994, Collins 1997). C'est une interprétation différente qui est défendue ici.
56. L'idiosyncrasie reste ainsi dans son lieu naturel : le vocabulaire, les mots tels qu'on les emploie.
57. On peut du reste envisager que le temps et le cas structural soient deux réalisations de la même catégorie I(nflection) (cf. Pesetzky & Torrego 2001 ; Nash 2002).
58. D'autres hypothèses sont certainement possibles, mais je ne m'attarde pas sur  $v$ , qui n'est pas mon objet.
59. Dans les langues sans classification comme le turc, en revanche,  $\{n\}$  est singleton.
60. Je ne parlerai ici que des noms (et donc aussi des adjectifs qui ne s'en distinguent pas en arabe, si ce n'est qu'avec eux, la valeur de  $\{n\}$  est une variable fixée par accord). Je ne prends donc aucun parti s'agissant des adverbes, des particules, etc.
61. Il s'ensuit que les préfixes de classe des langues bantu, par exemple, doivent être considérés comme des éléments radicaux et que, d'une façon générale, la préfixation est un type de composition (cf. Kihm à paraître (a)).
62. Jusqu'à un certain point, la fonction addition équivaut à l'opération dite *merger* de la DM. De même, la fonction inclusion dont il va être question équivaut à l'opération dite *fusion*.
63. On notera dès maintenant que toute la théorie des PI défendue ici est contenue dans (25). Je suis conscient que l'inclusion, interprétée comme je le fais, paraît exclure *a priori* les morphèmes-objets discontinus ou suprasegmentaux (tonals, par exemple). Je n'espère pas être en mesure de m'en justifier dans les limites de cet article.
64. Autrement dit,  $\{n_f\}$  est additionné ou inclus. Voir ci-dessous pour  $\{n\}$  associé aux masculins.
65. Cette affirmation vaut aussi des emplois adverbiaux de l'accusatif (*naSb*), qui sont clairement liés à l'insertion du lexème dans une certaine structure syntaxique.
66. Pour les collectifs tels que *šajarun* 'des arbres', on peut avoir recours à l'inclusion  $\{1 \{3 + \}\}$ , qui signifie que ces lexèmes sont à la fois singuliers et pluriels, mais singulier est l'élément dominant en tant que domaine de la fonction. Le singulatif correspondant, *šajaratun* 'un arbre', se forme alors aisément par suppression de  $\{3 + \}$ . Il y faut évidemment une notion du type "paradigme".
67. Cette affirmation, qui me paraît juste, a évidemment besoin d'une démonstration, que je ne peux entreprendre ici (cf. Corriente 1971 : 57). Je laisse aussi de côté le fait que

l'adjectif épithète d'un PI dénotant un humain mâle prend lui-même, autant que possible, la forme d'un PI (cf. *rijaalun Tiwaalun* 'des hommes grands' vs *rajulun Tawiilun* 'un homme grand'). Il n'est pas certain qu'on ait affaire là à un accord au sens strict du terme. (On n'oubliera pas que les adjectifs sont peu distincts des noms en arabe.)

68. Sauf, en arabe moderne, si le PE dénote une femme (cf. *malikaatun jamiilaatun* 'de belles reines' – mais *malikaatun jamiilatun* est possible).

69. On notera que la réanalyse de la consonne du préfixe comme  $C_1$  radicale s'accompagne d'un transfert du sens « Instrument » au gabarit global miCCaaC. On mesure aussi à quel point le terme de « relexicalisation » est mal choisi : *mu'allimun* n'est pas moins « lexical » que *miiftaaHun*. Je l'ai néanmoins conservé par commodité.

70. Cela étant, il faudra répartir très précisément les rôles respectifs du site post-C1 et du site post-C2. En effet, le PIU n'exclut pas que le site post-C1 soit parallèlement activé, mais cette activation est hors de son champ de vision, si l'on peut dire (cf. Kihm 2003 sur la formation des noms verbaux).

71. Je note les segments comme des valeurs de C et de V. C et V non valuées signifient que la position n'est pas phonologiquement interprétée.  $V_{\square}$  en (32) veut dire que V a une valeur, mais celle-ci n'est pas déterminée à cette étape.

72.  $\{v_{Tperf}\} = \{v \{T \{perf\}\}\}$  comme  $\{n_{pl}\} = \{n \{pl\}\}$ . Je suppose que {T} (temps) et {perf} (perfectif) sont des éléments lexicaux, sans prendre parti quand à leur éventuelle projection en syntaxe (voir plus haut).

73. Ce serait le cas si la forme était supplétive, par exemple *fus* par rapport à *suis*.

74. On verra bientôt pourquoi je n'adopte pas l'hypothèse autosegmentale de propagation d'une mélodie vocalique /a/ aux positions accessibles (cf. McCarthy 1981). V finale est non valuée parce que sa valeur /a/, qui peut ne pas être réalisée, est épenthétique, c'est-à-dire donnée par un processus phonologique ultérieur.

75. La supplétion totale (*fus* vs *suis*) se définit à présent comme l'application de la fonction d'épel non dédoublée à un ensemble d'éléments formé par inclusion.

76. Je remercie Nisrine Al Zahre (c.p.), qui m'a mis sur la piste de cette analyse.

77. Sauf à supposer que seule la première voyelle non spécifiée est phonologiquement interprétée, peut-être parce qu'elle est la seule à être réellement épelée. Mais cela apparaît comme une complication inutile, d'autant qu'il faudrait alors traiter de même les ségolés hébreux comme *kelev* 'chien', et expliquer pourquoi les deux premières voyelles y sont interprétées. Dans mon cadre, je suppose plus simplement que, pour les ségolés (et les mots arabes du type de *jabalun* 'montagne')  $f_V$  produit une suite  $V_{\square} V_{\square}$  interprétée une fois par la phonologie, donc avec même valeur pour les deux V.

78. Notez qu'il ne peut exister en arabe de paires minimales comme *cale* vs *colle*.

79. Il se peut que la valeur de C/V, /A/, /U/ ou /I/, soit non spécifiée à ce niveau, tout comme celle de la voyelle du singulier. Pour la clarté de l'exposé, je fais comme si elle était spécifiée.

80. Comme on l'a vu, la borne inférieure est toujours 3, mais  $C_3$  peut ne pas être valuée quand le site est inactif. Je continue à supposer qu'il n'existe pas de racines 1C.

81. Comme je l'ai indiqué plus haut, la valeur /a/ de  $V_1$  reste sans explication.

82. La fricative pharyngale /ʕ/ est bien du type glide en général, mais (a) elle n'est pas une réalisation possible de /A/ en arabe; (b) ce semble être un principe général que  $C_1$  n'est jamais affectée.

83. Le lecteur aura sûrement déjà remarqué que cette hypothèse s'accorde avec l'orthographe arabe et son emploi des *matres lectionis*. Pour autant que les solutions orthographiques reflètent les intuitions de locuteurs natifs dotés d'une sensibilité

linguistique particulière, pour autant, en d'autres termes, qu'une écriture est toujours aussi une analyse, j'y vois un autre point fort de la présente théorie.

84. En particulier les phénomènes de polarité et la valeur de  $V_1$  dans les PI du type *ġurfatun* / *ġurafun*.

85. Il y a des exceptions, assez rares. Lorsque le nombre de mores est le même, c'est le rythme qui diffère (cf. Kihm à paraître b).

## RÉFÉRENCES

- AIKHENVALD, Alexandra Y. (2000). *Classifiers: A Typology of Noun Categorization Devices*. Oxford: Oxford University Press.
- ANDERSON, Stephen R. (1992). *A-Morphous Morphology*. Cambridge: Cambridge University Press.
- ANGOUIARD, Jean-Pierre & M. DENAIS (1989). Le pluriel brisé en tigrigna, *Langues orientales anciennes, Philologie et Linguistique* 2: 99-148.
- ARAD, Maya. Why syntax matters, ce volume.
- ARONOFF, Mark (1994). *Morphology by Itself: Stems and Inflectional Classes*. Cambridge (Mass.): MIT Press.
- ASFOUR, Fouad-Martin (2001). Theoretische Aspekte der internen Pluralbildung im Klassischen Arabisch, Diplomarbeit zur Erlangung des Magistergrades, Universität Wien.
- BARTH, J. (1904). Beiträge zur Pluralbildung des Semitischen, *Zeitschrift der deutschen morgenländischen Gesellschaft* 58: 431-446.
- BEARD, Robert (1995). *Lexeme-Morpheme Base Morphology: A General Theory of Inflection and Word Formation*. Albany: State University of New York Press.
- BLACHÈRE R. & M. GAUDEFROY-DEMOMBYNES (1975). *Grammaire de l'arabe classique*. Paris: Maisonneuve & Larose.
- BOHAS, Georges (1997). *Matrices, étymons, racines: éléments d'une théorie lexicologique du vocabulaire arabe*. Leuven – Paris: Peeters.
- BOHAS, Georges & Jean-Pierre GUILLAUME (1984). *Étude des théories des grammairiens arabes, I. Morphologie et phonologie*. Damas: Publications de l'Institut français de Damas.
- CHOMSKY, Noam (1999). Derivation by Phase. Unpublished ms., MIT.
- CHOMSKY, Noam & Morris HALLE (1968). *The Sound Pattern of English*. New York: Harper & Row (*Principes de phonologie générative*, trad. des chap. 1, 2, 7, 8 et 9 par P. Encrevé, Paris, Le Seuil).
- COLLINS, Chris (1997). *Local Economy*. Cambridge (Mass.): MIT Press.
- COOPER, William S. (1974). *Set Theory and Syntactic Description*. The Hague: Mouton.
- CORRIENTE, Federico (1971). *Problemática de la pluralidad en semítico: el plural fracto*. Madrid: Consejo superior de Investigaciones científicas.
- DELL, François (1985). *Les Règles et les sons: introduction à la phonologie générative*. Paris: Hermann.
- DERENBOURG, Hartwig (1867). Essai sur les formes de pluriel en arabe. *Journal Asiatique* 9: 425-524.
- DRESSLER, Wolfgang U. (1989). Prototypical Differences between Inflection and Derivation. *Zeitschrift für Phonologie, Sprachwissenschaft und Kommunikationsforschung* 42: 3-10.

- FLEISCH, Henri (1961). *Traité de philologie arabe, vol. 1 : Préliminaires, phonétique, morphologie nominale*. Beyrouth: Imprimerie catholique.
- GREENBERG, Joseph H. 1955. Internal a-Plurals in Afroasiatic (Hamito-Semitic). Dans J. Lucas (ed.), *Afrikanistische Studien, D. Westermann zum 80. Geburtstag gewidmet*: 198-204. Berlin: Akademie Verlag.
- GUERSSEL, Mohand & Jean LOWENSTAMM (1990). The derivational morphology of the Classical Arabic verb. Ms, Université du Québec à Montréal, Université Paris 7.
- GUERSSEL, Mohand & Jean LOWENSTAMM (1996). Ablaut in Classical Arabic Measure I active verbal forms. Dans J. Lecarme, J. Lowenstamm & U. Shlonsky (eds.), *Studies in Afroasiatic Grammar*: 123-134. The Hague: Holland Academic Graphics.
- HALLE, Moris (2000). Distributed Morphology: impoverishment and fission. Dans J. Lecarme, J. Lowenstamm & U. Shlonsky (eds.), *Research in Afroasiatic Grammar*: 125-149. Amsterdam: John Benjamins.
- HALLE, Moris & Alec MARANTZ (1993). Distributed Morphology and the pieces of inflection. Dans K. Hale & S.J. Keyser (eds.) *The view from Building 20: Essays in honor of Sylvain Bromberger*: 111-176. Cambridge (Mass.): MIT Press.
- HAMMOND, M. (1988). Templatic transfer in Arabic broken plurals. *Natural Language and Linguistic Theory* 6: 247-270.
- KAYE, Jonathan, Jean LOWENSTAMM & Jean-Roger VERGNAUD (1990). Constituent structure and government in phonology. *Phonology* 7: 193-231.
- KAYNE, RICHARD S. (1994). *The Antisymmetry of Syntax*. Cambridge (Mass.): MIT Press.
- KIHM, Alain (2003). Arabic internal plurals and verbal nouns: morphological operations without morphemes. Communication au 17<sup>e</sup> Congrès international des linguistes, Prague.
- KIHM, Alain (à paraître (a)). Noun classes, gender, and the lexicon-syntax-morphology interfaces. Dans G. Cinque & R.S. Kayne (eds.), *Handbook of Comparative Syntax*. Oxford: Oxford University Press.
- KIHM, Alain (à paraître (b)). Les pluriels internes de l'arabe: un essai d'anatomie. *Bulletin de la Société de Linguistique de Paris*.
- KRATZER, Angelika (1994). External arguments. Dans E. Benedicto & J. Runner (eds.), *Functional Projections*. University of Massachusetts (Amherst): Occasional Papers 17.
- LADEFOGED, Peter (1982). *A Course in Phonetics*. New York: Harcourt, Brace, Jovanovich.
- LEVY, Mary Mansnerus (1971). *The Plural of the Nouns in Modern Standard Arabic*. PhD Dissertation, University of Michigan.
- LOWENSTAMM, Jean (1996). CV as the only syllable type. Dans J. Durand & B. Laks (eds.), *Current Trends in Phonology: Models and Methods*: 419-442. Oxford: Oxford University Press.
- MARANTZ, Alec P. (1982). Re-reduplication. *Linguistic Inquiry* 13: 435-482.
- MARANTZ, Alec P. (1997). No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. Dans A. Dimitriadis *et al.* (eds.), *Proceedings of the 21<sup>st</sup> Penn Linguistics Colloquium. UPenn Working Papers in Linguistics*, Philadelphia: 201-225.
- MAZAUDON, Martine (1997). /a/-glide in Sino-Tibetan. Communication au 16<sup>e</sup> Congrès international des Linguistes, Paris.
- MCCARTHY, John J. (1981). A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* 12: 373-418.

- MCCARTHY, John J. & Alan S. Prince (1990). Foot and word in prosodic morphology: the Arabic broken plural. *Natural Language and Linguistic Theory* 8: 209-283.
- MURTONEN, A. (1964). *Broken Plurals: Origin and Development of the System*. Leiden: Brill.
- NASH, Léa (2002). Entre la flexion et le verbe: syntaxe, morphologie, acquisition. Thèse d'habilitation, Université Paris 7.
- NASH, Léa & Alain ROUVERET (1997). Proxy categories in phrase structure theory. Dans K. Kusumoto (ed.), *NELS 27*. University of Massachusetts (Amherst): 287-304.
- PARTEE, Barbara H., Alice TER MEULEN & Robert E. WALL (1990). *Mathematical Methods in Linguistics*. Dordrecht: Kluwer.
- PESETSKY, David & Esther TORREGO (2001). T-to-C movement: causes and consequences. Dans M. Kenstowicz (ed.), *Ken Hale: A Life in Language*. Cambridge (Mass.): MIT Press.
- PETRÁČEK, Karel (1960). Die innere Flexion in den semitischen Sprachen: Entstehung und Entwicklung des Systems. *Archiv Orientální* 28: 547-606, 29: 513-545, 30: 361-408, 31: 577-624, 32: 185-222.
- PRUNET, Jean-François (1996). Guttural vowels. Dans G. Hudson (ed.), *Essays on Gurage Language and Culture: Dedicated to Wolf Leslau on the occasion of his 90<sup>th</sup> birthday, November 14<sup>th</sup> 1996*. Harrassowitz Verlag.
- PRUNET, Jean-François (1998). When vowels function like consonants. *Phonological Studies 1*, The Phonological Society of Japan, Tokyo Kataikusha: 219-226.
- RAFFELSIEFEN, Renate (1992). A non configurational approach to morphology. Dans M. Aronoff (ed.), *Morphology Now*: 133-162. Albany: State University of New York Press.
- RATCLIFFE, Robert R. (1998). *The 'Broken' Plural Problem in Arabic and Comparative Semitic: Allomorphy and Analogy in Non-Concatenative Morphology*. Amsterdam: John Benjamins.
- RITTER, Elizabeth (1991). Two Functional Categories in Noun Phrases: Evidence from Modern Hebrew. Dans S.D. Rothstein (ed.), *Perspectives on Phrase Structure: Heads and Licensing*. Vol. 25 of *Syntax and Semantics*. Academic Press, New York: 37-62.
- SAUSSURE, Ferdinand de (réed. 2002). *Écrits de linguistique générale*. Paris: Gallimard.
- SCALISE, Sergio (1984). *Generative Morphology*. Dordrecht: Foris.
- SÉGÉRAL, Philippe (1995). Une théorie généralisée de l'apophonie. Thèse de Doctorat, Université Paris 7.
- SPENCER, Andrew (1991). *Morphological Theory: An Introduction to Word Structure in Generative Grammar*. Oxford: Basil Blackwell.
- STUMP, Gregory T. (1990). Breton Inflection and the Split Morphology Hypothesis. Dans R. Hendrick (ed.), *Syntax and Semantics 23: The Syntax of the Modern Celtic Languages*, 97-119. New York: Academic Press.
- STUMP, Gregory T. (2001). *Inflectional Morphology: A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.
- WRIGHT, W. (1896-1898/1991). *A Grammar of the Arabic Language* (3rd edition). Cambridge: Cambridge University Press.

**ABSTRACT**

Arabic internal plurals of nouns and adjectives (also known as “broken” plurals) raise a challenge for morphological theory in that they do not obviously involve affixation of a plural morpheme analogous to, *e.g.*, English /-ez/. Rather they show a contrast of forms that share the same consonantal root, but differ in vocalization and possibly on other counts as well.

This article proposes an analysis of internal plural formation that derives all forms through a single process, *viz.* inserting one of the three glides /I/, /U/, and /A/ into a specific *site* located between the 2nd and 3rd root consonants, the post-C<sub>2</sub> site, thereby creating an onset (C) or a nucleus (V) position.

It then explores the consequences of this account for the overall organization of grammar. A modified version of Distributed Morphology is put forward that construes the lexicon as a set of sets of elements endowed with meanings but no form, and where Spell Out is a *function* that takes element sets and returns legitimate phonological objects, *i.e.* CV strings.

**KEYWORDS**

CV phonology, functions, glides, (non) concatenative morphology, set theory, site, template.